

Please send to Applicants
Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 15:12:48 ; Search time 590 Seconds
(without alignments)
15500.940 Million cell updates/sec

Title: US-10-030-850-1
Sequence: 1 gatctgcagcagccgctc.....ggatcaaatcataaaaaa 5145

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued Patents_NA:*
2: /cgn2_6/prodata/1/ina/1/COMB.seq:*
3: /cgn2_6/prodata/1/ina/5/COMB.seq:*
4: /cgn2_6/prodata/1/ina/6/COMB.seq:*
5: /cgn2_6/prodata/1/ina/8/COMB.seq:*
6: /cgn2_6/prodata/1/ina/10/COMB.seq:*
7: /cgn2_6/prodata/1/ina/12/COMB.seq:*
8: /cgn2_6/prodata/1/ina/14/COMB.seq:*
9: /cgn2_6/prodata/1/ina/16/COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	732.8	14.2	767	US-09-270-767-596	Sequence 596, App
2	732.8	14.2	767	US-09-270-767-15878	Sequence 15878, A
3	661	12.8	1420	US-09-270-767-6018	Sequence 6018, A
4	661	12.8	1420	US-09-270-767-21300	Sequence 21300, A
5	452.4	8.8	466	US-09-270-767-14266	Sequence 14266, A
6	149.8	2.9	3153	US-10-104-047-327	Sequence 327, App
7	149.4	2.9	154	US-09-270-767-30401	Sequence 30401, A
8	141.8	2.8	2524	US-10-104-047-517	Sequence 517, App
9	116.6	2.3	1906	US-10-104-047-113	Sequence 113, App
10	116.6	2.3	6445	US-09-620-312D-324	Sequence 324, App
11	116.6	2.3	6487	US-09-620-312D-323	Sequence 323, App
12	47.2	0.9	7218	US-08-232-463-14	Sequence 14, App1
13	43.8	0.9	7218	US-08-232-463-14	Sequence 14, App1
14	42.8	0.8	3472	US-08-232-463-14	Sequence 14, App1
15	41.6	0.8	3684	US-09-561-077C-25	Sequence 25, App1
16	41.6	0.8	3684	US-09-561-077C-25	Sequence 25, App1
17	41.6	0.8	7113	US-09-561-077C-26	Sequence 26, App1
18	41.6	0.8	7113	US-09-561-077C-26	Sequence 26, App1
19	41.2	0.8	1713	US-09-252-991A-10767	Sequence 10767, A
20	41.2	0.8	2490	US-09-252-991A-10579	Sequence 10579, A
21	41.2	0.8	3384	US-09-252-991A-10278	Sequence 10278, A
22	41.2	0.8	19124	US-08-487-826B-13	Sequence 13, App1
23	41.2	0.8	154746	US-09-827-688-8	Sequence 8, App1
24	40.2	0.8	468	US-09-252-991A-2152	Sequence 2152, App

C	25	40.2	0.8	777	3	US-09-252-991A-1732	Sequence 1732, Ap
C	26	40.2	0.8	930	3	US-09-252-991A-1997	Sequence 1997, Ap
C	27	40	0.8	1290	3	US-09-902-540-5006	Sequence 5006, Ap
C	28	40	0.8	12950	3	US-09-902-540-1036	Sequence 1036, Ap
C	29	38.8	0.8	505	3	US-09-621-976-15639	Sequence 15639, A
C	30	38.6	0.8	2358	3	US-09-252-991A-16218	Sequence 16218, A
C	31	38.6	0.8	2358	3	US-09-252-991A-16218	Sequence 16218, A
C	32	38.4	0.7	1026	3	US-09-270-767-12503	Sequence 12503, A
C	33	38.4	0.7	3395	3	US-09-902-540-3658	Sequence 3658, Ap
C	34	38.4	0.7	23847	3	US-09-902-540-1177	Sequence 1177, Ap
C	35	38.2	0.7	390416	3	US-09-949-016-16923	Sequence 16923, A
C	36	37.8	0.7	411	3	US-09-252-991A-16270	Sequence 16270, A
C	37	37.6	0.7	996	2	US-08-924-759-21	Sequence 21, App1
C	38	37.6	0.7	996	3	US-09-248-335-21	Sequence 21, App1
C	39	37.6	0.7	2754	3	US-09-489-039A-5411	Sequence 5411, Ap
C	40	37.4	0.7	1652	3	US-08-726-214-17	Sequence 17, App1
C	41	37.4	0.7	3639	3	US-09-902-540-5646	Sequence 5646, Ap
C	42	37.4	0.7	3978	3	US-08-726-214-1	Sequence 1, App1
C	43	37.4	0.7	51354	3	US-09-902-540-1270	Sequence 1270, Ap
C	44	37.2	0.7	50000	3	US-09-662-254B-26	Sequence 26, App1
C	45	37.2	0.7	160759	3	US-09-949-016-16514	Sequence 16514, A

ALIGNMENTS

QY	2639	CCTGATCAGCCTTACACAGTATCTGGCCGATGGAAGAGCTTACTCAGAGCACTTAA	2639	US-09-270-767-596	Sequence 596, Application US/09270767
DB	1	CCTGATCAGCCTTACACAGTATCTGGCCGATGGAAGAGCTTACTCAGAGCACTTAA	60	Patent No. 6703491	
QY	2699	ATGCGTTCGACATGTTCCGAGATGTTCAATGATACACACCAAGGTACTACT	2758	GENERAL INFORMATION:	
DB	61	ATGCGTTCGACATGTTCCGAGATGTTCAATGATACACACCAAGGTACTACT	120	APPLICANT: Homburger et al.	
QY	2759	GTTTTGGACAGCGTACCGAATCTGATTTGTTTCAACCAATGAGCCGATCTACTATTCGAT	2818	TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster	
DB	121	GTTTTGGACAGCGTACCGAATCTGATTTGTTTCAACCAATGAGCCGATCTACTATTCGAT	178	FILE REFERENCE: File Reference: 7326-094	
QY	2819	CTCTTCATCTTTGTTGAGCGGTTGTTTCAACCAATGAGCCGATCTACTATTCGAT	2818	CURRENT APPLICATION NUMBER: US/09/270,767	
DB	179	CTCTTCATCTTTGTTGAGCGGTTGTTTCAACCAATGAGCCGATCTACTATTCGAT	237	CURRENT FILING DATE: 1999-03-17	
QY	2879	GCAACAGCAATTTGGAAGAGCTTACAGGTGTCAGATATTTCCCGAGGATTTGCA	2938	NUMBER OF SEQ ID NOS: 62517	
DB	238	GCAACAGCAATTTGGAAGAGCTTACAGGTGTCAGATATTTCCCGAGGATTTGCA	297	SOFTWARE: PatentIn Ver. 2.0	
QY	2939	GCTTAAAGAACTATTAGATCATATGATTTCTACTCAACCCCACTTCAAGAGCGC	2998	SEQ ID NO 596	
DB	298	GCTTAAAGAACTATTAGATCATATGATTTCTACTCAACCCCACTTCAAGAGCGC	357	LENGTH: 767	
QY	2999	CATGCCATTTCGCGCTTCTTACGATCTGCGCAATACATCTGTAAGACGCACTT	3058	TYPE: DNA	
				ORGANISM: Drosophila melanogaster	

Query Match 14.2%; Score 732.8; DB 3; Length 767;
Best Local Similarity 98.7%; Pred. No. 8e-208;
Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;

Db	358	CATGCCCATTTCTGGCCTTCTTTACGGATCTGGCCAAATACATATCTTTAAAGCAGGACGATT	417
OY	3059	TCCGACGACACAGAGACCAGGCGCTGGAGCGGACGGTCTTAAAGCTGGCCGACGTGGCGGC	3118
Db	418	TCCGACGACACAGAGCCAGGCGCTGGAGCGGACGGTCTTAAAGCTGGCCGACGTGGCGGC	477
OY	3119	GGATCAGAAGTGATGATGAGTGGCTGCCGAGCCAGAGCTGCTAGTGGCCCTGTATPA	3118
Db	478	GGATCAGAAGTGATGATGAGTGGCTGCCGAGCCAGAGCTGCTAGTGGCCCTGTATPA	537
OY	3179	CCTGAATACCCCGCAATGACCCCTTTTACTGGCCGACCTGCCAAAGGTATATCAGGACTC	3238
Db	538	CCTGAATACCCCGCAATGACCCCTTTTACTGGCCGACCTGCCAAAGGTATATCAGGACTC	597
OY	3239	TGCCCGATCCCTGACATCCCATTTGGCAATAGAGGGGGGAAAGCCAAATTTGCATTTGGGGTGC	3298
Db	598	TGCCCGATCCCTGACATCCCATTTGGCAATAGAGGGGGGAAAGCCAAATTTGCATTTGGGGTGC	657
OY	3299	CAATTGCGCTGTACTCTCCATTGAGAGATGACGATCCCAAGCTTTGGCAAAGTCCCTC	3358
Db	658	CAATTGCGCTGTACTCTCCATTGAGAGATGACGATCCCAAGCTTTGGCAAAGTCCCTC	717
OY	3359	TGTGGGATCCATTTGGCTGCTTCAGAGCCACACACCACTTACATCA	3408
Db	718	TGTGGGATCCATTTGGCTGCTTCAGAGCCACACACCACTTACATCA	767

```

RESULT 2
US-09-270-767-15878
Sequence 15878, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 15878
LENGTH: 767
TYPE: DNA
ORGANISM: Drosophila melanogaster
US-09-270-767-15878

```

Query Match	Similarity	14.2%	Score 732.8	DB 3	Length 767
Best Local	Similarity	98.7%	Prod. No. 8e-208		
Matches	760	Conservative	0	Mismatches	7
				Indels	3
				Gaps	2
Qy	2639	CCTGATCAGCCTCACACAGTATCTGGCCGATGGAAAGAGCTCACTCAGACCACTTAA	2698		
Dh	1	CCTGATCAGCCTCACACAGTATCTGGCCGATGGAAAGAGCTCACTCAGACCACTTAA	60		
Qy	2699	ATGCGTTCTGACATGTTCCGACAGATGTTCAATGATACACACACCAAGGTGTACTACT	2758		
Dh	61	ATGCGTTCTGACATGTTCCGACAGATGTTCAATGATACACACACCAAGGTGTACTACT	120		
Qy	2759	GTTTTTGGACACGCGTACCGAATCTGATTCATACGAAACGAAACTTCACGAAATNG	2818		
Dh	121	GTTTTTGGACACGCGTACCGAATCTGATTCATACGAAACGAAACTTCACGAAATNG	178		
Qy	2819	CTTTCATCTTGTGTGACGCGGTTTCAACCAATTTGGGACCCGATCTACTCAATTCCGAT	2878		
Dh	179	CTTTCATCTTGTGTGACGCGGTTTCAACCAATTTGGGACCCGATCTACTCAATTCCGAT	237		
Qy	2879	GCAACGACGAATTTGGAAGACGCTACAGGTGTCCACGAATATTTTCCCAACGCGAGTTTCA	2938		
Dh	238	GCAACGACGAATTTGGAAGACGCTACAGGTGTCCACGAATATTTTCCCAACGCGAGTTTCA	297		
Qy	2939	GCTTAAACAACTATTAAATCATATTCGGATTCTAATCAACCCCAATCAACCAAGACGCG	2998		
Dh	298	GCTTAAACAACTATTAAATCATATTCGGATTCTAATCAACCCCAATCAACCAAGACGCG	357		

QY	2999	CATGCCATTCTGGGCTTCCTTACGAGTCTGGCCAAATACATCTGTAAAGACAGACTT	3058
Db	358	CATGCCATTCTGGCTTCCTTACGAGTCTGGCCAAATACATCTGTAAAGACAGACTT	417
QY	3059	TCCGAGCAGACAGAGCCAGGCTTGCGAGCGGACCGTCTTAAAGCTTGCCCACTGGCGGC	3118
Db	418	TCCAGCGACAGAGCCAGGCTTGCGAGCGGACCGTCTTAAAGCTTGCCCACTGGCGGC	477
QY	3119	GGATCAGAAGTCAGTGAAGCTGGCGCTCCAGGCCAGAGAGCTGCTTAGTGCCCTGTATTA	3178
Db	478	GGATCAGAAGTCAGTGAAGCTGGCGCTCCAGGCCAGAGAGCTGCTTAGTGCCCTGTATTA	537
QY	3179	CCTGAATAACCCCGGAAATGACCCCTTTTAACTGGCCGACCTGCGAAAGGATATCAAGACTC	3238
Db	538	CCTGAATAACCCCGGAAATGACCCCTTTTAACTGGCCGACCTGCGAAAGGATATCAAGACTC	597
QY	3239	TGCCGATTCCTGCATCTCATTTGCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGC	3298
Db	598	TGCCGATTCCTGCATCTCATTTGCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGC	657
QY	3299	CAATTGCGCTTAGTAGCTCTCTCCATTGAGCAAGTAGCAGTTCCTCAAGCTTTGCAAAGTCCCTC	3358
Db	658	CAATTGCGCTTAGTAGCTCTCTCCATTGAGCAAGTAGCAGTTCCTCAAGCTTTGCAAAGTCCCTC	717
QY	3359	TGTGGGTCCTATTTCCTCGCTTACAGAGCCACACCAACCTTAGCATCA	3408
Db	718	TGTGGGTCCTATTTCCTCGCTTACAGAGCCACACCAACCTTAGCATCA	767

```

RESULT 3
US-09-270-767-6018
; Sequence 6018, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6018
; LENGTH: 1420
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-6018

```

[illegible]

Please send to Applicants
Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 14:59:23 ; Search time 54 Seconds
(without alignments)

2658.434 Million cell updates/sec

Title: US-10-030-850-2

Perfect score: 7611
Sequence: 1 MAYRRPSDDGFIQWPKAD.....RNCISGGSSSTKNSAASS 1492

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 80: *
2: PIR: *
3: PIR: *
4: PIR: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1375	18.1	1289	T00387	KIAA0622 protein -
2	1348	17.7	1324	T00386	hypothetical prote
3	611	8.0	1063	E88546	protein R107.6 (im
4	593.5	7.8	1080	S30876	hypothetical prote
5	389	5.1	826	C88550	protein ZC84.3 (im
6	328	4.3	1972	S68176	TOG protein - huma
7	277.5	3.6	730	S28294	hypothetical prote
8	262.5	3.4	643	S28293	hypothetical prote
9	259.5	3.4	1462	T11658	probable mitotic s
10	198.5	2.6	5327	T13564	microtubule-associ
11	198	2.6	1148	C86403	probable transposo
12	195	2.6	199	F84586	hypothetical prote
13	195	2.6	2271	F90073	hypothetical prote
14	192	2.5	2232	T34434	hypothetical prote
15	192	2.5	4377	A55575	ankyrin 3, long sp
16	190	2.5	860	T39502	hypothetical prote
17	184	2.4	1032	T34433	hypothetical prote
18	182.5	2.4	1955	T30934	myosin-like protei
19	180.5	2.4	1630	A53577	ascies ataloglyco
20	180	2.4	809	T41645	probable spindle p
21	179.5	2.4	1513	T50073	mitotic spindle pr
22	179	2.4	1727	T50073	myosin-like coiled
23	177.5	2.3	3507	T34513	hypothetical prote
24	177	2.3	1108	T17455	translation initia
25	177	2.3	1696	T00057	hypothetical prote
26	175.5	2.3	2346	T13829	tyr homolog - frui
27	174.5	2.3	2021	A84771	similar to ch-TOG
28	173	2.3	1415	T21241	zys-9 protein - Ca
29	172.5	2.3	1114	T14351	serine/threonine-s

30	172	2.3	725	2	A41258	a-agglutinin core
31	172	2.3	1104	2	S59310	probable membrane
32	172	2.3	3225	2	I52300	giantin - human
33	172	2.3	3259	1	A56539	giantin - human
34	171	2.2	927	2	T38127	phosphoprotein - f
35	171	2.2	1326	2	AC3372	kinasin-like prote
36	170.5	2.2	2481	2	D90011	Fmb protein (limpo
37	170	2.2	1792	2	T20363	hypothetical prote
38	169.5	2.2	6713	2	B89921	hypothetical prote
39	168	2.2	952	2	T18837	hypothetical prote
40	168	2.2	1912	2	T29088	vitellogenin I pre
41	167.5	2.2	1520	2	T44231	hypothetical prote
42	167	2.2	1237	2	T45070	protein kinase hom
43	167	2.2	2022	2	T43214	ovl protein - nem
44	164	2.2	2819	2	A90551	conserved hypotet
45	163.5	2.1	1072	2	A86827	hypothetical prote

ALIGNMENTS

RESULT 1

KIAA0622 protein - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C/Accession: T00387

R/ Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, DNA Res. 5, 169-176, 1998

A/Title: Prediction of the coding sequences of unidentified human genes. X. The comple

A/Reference number: Z14142; MUID:98403880; PMID:9734811

A/Accession: T00387

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1289 <IS>

A/Cross-references: UNIPROT:O75118; UNIPARC:UP1000007259A; EMBL:AB014522; NID:93327057

A/Genes:

A/Note: KIAA0622

Query Match	18.1%	Score 1375;	DB 2;	Length 1289;
Best Local Similarity	30.1%	Pred. No. 5,7e-69;		
Matches 416;	Conservative 235;	Mismatches 495;	Indels 234;	Gap 46;
QY	236	PSAKNTNGVGLDEADNIGLRERPTMKPLHSAVSSLRPKPNVDVTGDAGATM	295	
DB	15	PSRRY-----VGMGTRRLG-----	50	
QY	296	ESFESSFEVVPQNTFPHADMDYKQVVIISDKNADMEKVDALKTRALLISYHQ	355	
DB	51	EDFIKAFDDVPVQVYSSRDLESINKIREIISDDKGDWQRYNALKTRISLLAGAAEY	110	
QY	356	PQVAVQLELSFVDILKELEISQVIREACITTAVMKTLRNKLDACFMSILEHINL	415	
DB	111	DNFQO-HLELLDGA-F-KLSAKDLRSQVVEACTTGLHLSVLANKFDHGAEMIPITFNL	168	
QY	416	IQNSAKVIASASTIALKVIKTYHAFLKIYDTLQNSKSDIRSTLCMLVLFEEWQ	475	
DB	169	IPNSAKIMATSGVAVRLIRHTHPRLPVITSNT-SKSAVARRCEFDLLQEWQ	227	
QY	476	TYALEFNATVLDTLKKSIGDADCDARRSRAYAFRRHFPBLADQITGLDIAQRL	535	
DB	228	TSLSERHISVLAETIKGICHADSEARIFARCKYGFHSHFRFAHLVHTLESSYQAL	287	
QY	536	ERREBEGGGGGTGCTGTAPEPTRTVSRIGTPGTQKTPMSRSISAVDTAAQAQYA	595	
DB	288	QSHLK-----NSDSIVSLPQSDRSS-----SSQESINRLSAKRSFTGTSASIVSTK	338	
QY	596	AQYTLVSRQKPLGPNNSQAASMTGAASGLPRPLNSNGTP-----ATPGSV---	647	
DB	339	SVSTISLQK-----SRSDIDVNAASA-----SKSVSSSSGTTTPSSAAALPPGSYASL	388	
QY	648	-----TPRPRGAG-----VQSO-----PSRSRSTPSTKLRDQY	677	

Db 389 GRIIRTRSGSGANVASTPDRNGRGRRAKVVSGQSRSGANPAGAGSRSSSPGKLIGSGY 448
 Qy 678 GGI-GNYVRG-ATGAI PKKASGI PRSTASRETSPTRSQ---GGLMKRMYSTGAG---S 729
 Db 449 GGITGGSSSGPPVTPSESEKRSKI PKQSGSRETSPPRIGLASSRIPRSMGCGSRDTS 508
 Qy 730 RRTPERNNFVRSPAPARLLAOSREAHETLGVGDQDQ-PPVSGDYNRSGG----- 778
 Db 509 RSSNDTSPARGPPL-----DRGLGQPGRI PGSVAMRVLSTLDLEAADA 558
 Qy 779 ---MWRKLMGRDESDDI DEASAVGSEBSPDSYTTGNKSNVSLSGSHRLDMSTOR 834
 Db 559 LKKPVRRRYEPYGMVSDDDANSPASSVCSERSYGS--RNGGIPHY-----L 602
 Qy 835 APFDIETIIIOFCASHTWSEBKDGLSLNOYADGKELQOOLKCYLDMFRKKPMOTHTK 894
 Db 603 ROTEDVAEVLNHCASSNWERKEGLGLONLKSQRTLSRVELKICEIFTRFADPHSK 662
 Qy 895 -VYSLPLDTVTTELILVHANETSRNGSSCLTRLFNKLTGDLNLSMHSKIMKTLOVHVEYF 953
 Db 663 RYFSMFLTELVTPELIIHKDL-QDMLFVLLTQLKMGADLLGSVQAKVQKALDVTBDSF 721
 Qy 954 PPOLOLKELFRIISDSTOTPTTKTRAIIRFLTDLANTYCKSSDPSPD--QSOACERTVL 1011
 Db 722 PPDQOFNIMRFIVDQOTQTPNKLKVKVAILIKYTESLA---RQWD-PTDFVNSSETRLAVS 776
 Qy 1012 KLAQLAADQKSMELRSGARCLVALNLTNPQWTLTLADLPKYQOASARCSJSHMRQS 1071
 Db 777 RIITWTTEPKSSDVRCAAQIVLISLELMTPEFTMLGALPKTPQDGAIKLILHNLKNS 836
 Qy 1072 QSCNSGANSPPS-----SPLSSSPKPLQSPS-----VGPPA-----SLQSHHOL 1112
 Db 837 ---NNSVGSPTMTIGRTPSRHTSSRTSPILSPNCSHGGLSPRLMGWAGADLAKPRPF 893
 Qy 1113 STSGTSPPRSQSSVEOELLFSSSELDIQ-HNIQKTSSEIHCRCGQYQTLAPNGRHLIQ 1171
 Db 894 SQPNSIPRTAPSHKALRSYSPMLDYDTEML--NSEIIVSYSLRGVTE-AIEKFSFSGQSD 950
 Qy 1172 YHDOGGOD---SCASLSNSKTQSSANTTQSNTP--ESATMRLLDNLERETTONAKS--- 1223
 Db 951 LNEPIKRDGKKECDIVSRGGAASPAKTEGGSGEVEGRTALDNKTSILANTOPPRAPFPP 1010
 Qy 1224 -----PTDDAKVITVSINNAENGEL-----ILASNMISESV 1254
 Db 1011 RADVNPVYPSDA-----INTYDKTALKEAVFDDMEQLRDPVIDHSDLVADLKELESN 1064
 Qy 1255 VAVALTLLTQDQVELLQ-TSLTNLGCICIGANCELPNKFPRSLMRLLMLILAEHTDVI 1313
 Db 1065 HNERVEERKGALELLKTRIEDSLGW-----EEHFKTILLLLETLTDKCHSIRA 1115
 Qy 1314 AGHLVLSKIMRSNKNRMHMFLELLKTIQCYGHS-KEALRDIDSMIPRIAPSLPLDL 1372
 Db 1116 LALRVARELIRNOPAR--EKVYAEITIKTLEHAKSHKEVVAABEAASTLASSIHPEQ 1173
 Qy 1373 SINIVNVIATGEFTNLCAIKILLVETHEHSGEITDAHLDIVFPNLASADDTOSWAK 1432
 Db 1174 CIKVLCPITQTDADYPIINLAIKMQTKVVERIAKESLLQLLVIIIPELLGYONTBSSYRK 1233
 Qy 1433 AAVFCLVLYFVGEKVKPKSVLNPSTKRLINVIKORNCISGSGSTKNSAASGS 1492
 Db 1234 ASVFLVALIVSYIGED-LKPHLAQLTGSKKLLNLYIKR-----AQTNNSSSSSS 1283

RESULT 2

T00386
 hypothetical protein KIAA0627 - human (fragment)
 C1:Species: Homo sapiens (man)
 C1:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
 C1:Accession: T00386; T46487
 C1:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
 DNA Res. 5, 169-176, 1998
 A1:File: Prediction of the coding sequences of unidentified human genes. X. The complete
 A1:Reference number: Z14142; M0ID:98403880; PMID:9734811

A1:Accession: T00386
 A1:Status: preliminary; translated from GB/EMBL/DBJ
 A1:Molecule type: mRNA
 A1:Residues: 1-1324 <18>
 A1:Cross-references: UNIPROT:075122; UNIPARC:UPI000006f360; EMBL:AB014527; NID:G3327067
 A1:Experimental source: Dr. J. Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 A1:Submitted to the Protein Sequence Database, January 2000
 A1:Reference number: 223035
 A1:Accession: T46487
 A1:Status: preliminary
 A1:Molecule type: mRNA
 A1:Residues: 969-1324 <AAA>
 A1:Cross-references: UNIPARC:UPI000016ACCA; EMBL:AL137636
 A1:Experimental source: adult testis, clone DKFZp34F075
 A1:Genetics:
 A1:Note: KIAA0627; DKFZp34F075.1

Query Match 17.7%; Score 1348; DB 2; Length 1324;
 Best Local Similarity 29.6%; Pred. No. 2e-67;
 Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;

Qy 237 SALK-----NTNGNGVGLDEADNIGLRERPTMIKRPILHSAYSSSLRPKNVNDVTGDDGA 292
 Db 55 SAFKVPARKTSGN-----PANSARKP-----GSAGGPVVGASKEGAGGA 94
 Qy 293 VTMESEFSPFVWPOLNIFHAKMDMDIYKOVLIISDGNADMEKRVDAIKIRALLI--- 349
 Db 95 VDEDFIATFDVDSIQYSSRELEETLANKREILSDKHMDQANMLKIRLSLVAGA 154
 Qy 350 LSYHTQPOFVAVOLKEISLSEFVDILKEELRSQVIREACITIAVMSKTLNKLDAFCWSIL 409
 Db 155 AQYDFEFOHRLDLQALKLS-----AKDLRSQVAREACITIAHLSTVIGNKFHDGAEAIV 209
 Qy 410 EHLINLONSANVAVASASTIALKTYIKYTHAPKILKITYTDLNOSKSDISTICEMLVL 469
 Db 210 PTLFVLVNSAKVATSGCAAI RPIRTHVPRILPILTSNCT--SKSVVRRRSFEPIDL 268
 Qy 470 LFEENQATLRNNAITVLEDTLKSIGDADCDARHRSRYAVAFRRHPPELADQIYGTIDI 529
 Db 269 LLQEQHTSLERHAAVLVETIKKGIHDDADARVARKTYGRLNHFGEAETLYNSLEP 328
 Qy 530 AAQRALEEREREGGG-----GATGTGTAPETRTVRSRI 564
 Db 329 SYQSKLQYVLLKSSGVSASLPQSDRSSSSQESLNRPFSSKMTANPSTVAARVSGSSKA 388
 Qy 565 GRTPECTLOKTPPMSRISAVDTAAQRAKVAQVTLVYRQKRPGLPNNSQAS----- 617
 Db 389 SSLPESLD-----RSRSDIDIVNAAAGAKAH-HAAGQVRRGRUGAGALNAGSTASLEDT 441
 Qy 618 ---MTGAAGSGLPRLNNSNGTTPATTPGSVTPRPRGRAG---VSQOPGSRSTSP-- 669
 Db 442 SDKLDGTASEGRRRAKLSA-----PLAGMNAADSGRGRRTKVNVSQOGSRGSGR 496
 Qy 670 ---STKLDYQGGIGNYRGATGAI PKKASGIPSTASSTRETSPT---SGGGLMKRMY 723
 Db 457 VLTTLTALSTVSGVQVRLVNSASA--QKRSKIPSGGCSREASPSRLSVARSRRIPRSPV 554
 Qy 724 STGA---GSRTPERNNVPYRSP-----ARLAAQ 750
 Db 555 SQGGSREARSRESSDTSVRSFQPLASRHSRSTGALYAPVYAGSGGYIGISQSRSLSS 614
 Qy 751 SREAEHTLVGDDQPDVYSGDYNRSGGMGRKLMGRDESDDI DEASAVCSERSFPDS 810
 Db 615 SVSARVAVNTGSD--VEEAVADALKKPRARR--YESYGHSDDDANSDASCSERSYS- 670
 Qy 811 YTRGNKSNVSLSGSHTRLDMSTQAPPDITITIIQFCASHTWSEBKGLISLTQYLDGK 870
 Db 671 -----RNGSI-----PTYRKOTEDVAEVLNRCASSWSMSEKEDLLOMLKOR 715
 Qy 871 ELTQOOLKCVLDMFRKPMNDHTK--VYSLFDITYTELILVHANETSRNGSSCLTRLFNK 929
 Db 716 TLSRVELKRLCEIFTRMFADPHGRKVSFWLETLVFIQVHKDL-QDMLFVLLTQLLKK 774

Db 361 ATCGAAGTACTACTTACGAGGAGCCACCGCGCTTGACTGAGACAGCCAGCGGCC 420
 QY 2573 ATTGACGACATGAGACAGATATTTCAGTTCTGCCGATGAGCAT 2618
 Db 421 ATTGACGACATGAGACAGATATTTCAGTTCTGCCGATGAGCAT 466

RESULT 6

US-10-104-047-327
 Sequence 327, Application US/10104047
 Patent No. 6943241
 GENERAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: No. 6943241el full length cDNA
 FILE REFERENCE: H1-A0105
 CURRENT APPLICATION NUMBER: US/10/104,047
 CURRENT FILING DATE: 2002-03-25
 PRIOR FILING DATE:
 NUMBER OF SEQ ID NOS: 4096
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 327
 LENGTH: 3153
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-104-047-327

Query Match 2.9%; Score 149.8; DB 3; Length 3153;
 Best Local Similarity 47.1%; Pred. No. 2.9e-33;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

QY 232 CATGCTGACGCGGACGCCATTAAAGATTGCAACAAAAGTCCCTGGAGCGTTCTCGAGC 291
 Db 521 CTGGGTGAACCTTAGCAATTACAAAGTGTCTGCTGGGACATGACATCTGTCGCC 580
 QY 292 TAATTAAGCATTTGGGCGAGCCATTAAATGATACAGCGCTACCGTTCTGCCATGTGA 351
 Db 581 TGGTACCCTGGTGAAGATCGCTTCAAGCGCAATGGGCAATGCTGCTGCAAGTCTTA 640
 QY 352 TCGATCGCTGGAGACAGACAGACAGACAGATCCGCGAGAGCGCACTTCTGCGCGG 411
 Db 641 TAGACAGACTAGAGATGCTAAAGACTGTGAGGAGAGAGACCAAACTGCTGCTTAA 700
 QY 412 ACCTGAGGACAGAGTCTTCCGCCCAAGCGCTGATCGACAGTGGCCACTAGCT 471
 Db 701 AATGATGATCAAGCTGCTAATCCCACTAGCTATGAGACAGATGCT-----TGAG 754
 QY 472 GCTTCAAGACAAAGACGCAAGTGGCGAGAGTCTCTTCAAGCATTTGAGCGCTG 531
 Db 755 GCTTCAAGACAAAGATTTCCGATCTGAGAGAGCATCTGCTGCTTATGCAACAC 814
 QY 532 TTGATGATGCGGACCCAGCAAGCTTATGTTGCTCTATTAACCAACAGTTTGTGAC 591
 Db 815 TCAATGCTCTGAGACAGACACTTAACTAAGCAAGATTGTCACATATATGCACT 874
 QY 592 TTCTCGAGATCCCAAGTTAATGTGAGGAGCGGCATCCAAAGCTAGTGAATCT 651
 Db 875 TACTTGAGATCCAAAGCCAGGTTGAGATGAGCAATTAACAGTTTGAATTT 934
 QY 652 ACAAGCATGTAAGGAGATGATGCGCCAGACTCCCTGCGATGAGCATGTTCTGCT 711
 Db 935 ACAAGCATGTAAGGAGATGATGCGCCAGACTCCCTGCGATGAGCATGTTCTGCT 991
 QY 712 CGAATGCTATGTTGAGCAAAAGTTGACCGAGTCAACAGAGAGGTCTACTGCTAC 771
 Db 992 CCGGTTGATGATATTTTACAAATTTGATGAATCCAAATCTGGAACATGATAC 1051
 QY 772 CTGAGCCCTTAAACAGATGCAATGAGAGGCTTGACAGAGCC-GACATATT 830
 Db 1052 AATCGCAATGATTAATAATTTGACGATGAAGATTTCTGTGATGTAACAGACTTCT 1111
 QY 831 GGGTTGAGAGGAGCCAGACAGATGATTAAGCGGCACTAAGCTGCGCGTTTCTGCA 890

Db 1112 CTGCTAGTCTACATCAATCCAAAGGCTCACCAAGTTCTCGAGAAAGATTGGA 1171
 QY 891 -----TCACTGCGCCCAAAACCAATGTGAACGATGTGACCGGTATG 933
 Db 1172 CCACCCCGCGCTGTTGATTCACACCTTGATCCAAAGTCTTACGCTGCAAAAGAG 1231
 QY 934 CCGGCGCGTAAACAGTGAATCTTTGATATCTAGCTTTAGAGTGTCCGCAATTGAACA 993
 Db 1232 CTGCTGCTGTTGATGAAGATTTTAAAGCATTTGATGATGATCTGTTAGTACGA 1291
 QY 994 TCTTCACGCTAAGACATGAGATGATCTCAAGCAAGTACTAGTATCATCACTGATA 1053
 Db 1292 TTTATTCAGCGGACCTTGAGATCTCAATTAAGAAATTAAGAAATATATCTGATG 1351
 QY 1054 AAAAGCACTGAGAGAAACGTGTGATGCTCTCAAGATGTCAGGCGATTCATTC 1113
 Db 1352 ACAGCATGATGAGAGAGAGAAATGCTTAAAGATTTGATTTACTTTT 1409
 QY 1114 TCAGCTATCACTAGCGCGCATTTGCTGTACAGCTAAAGAAATGCTTAAAGCT 1173
 Db 1410 -----GGCTGTGCTGTGATGATGATTAATCTTTCAATTTGCTTTTGTGATGAG 1465
 QY 1174 TCGTGACATCTCAAGAGAGAACTAGATCAAGGTGATCCGCGAGCGTATCAACA 1233
 Db 1466 CTTTAACTCTGCTGTAAGAGACCTGCGCTCAGGATGAGGAGCGTGTATCAGCT 1525
 QY 1234 TCGCTACATGCTAAGCAAGCTGAGAAATTAAGTATGCTTCTGCTGAGCATTTTG 1293
 Db 1526 TGGGAGTCTGTATACAGTTCTGAGAAATTAAGTTGACATGAGCTGAAGCATTAATG 1585
 QY 1294 AGCACTGATTAATTAATACAGAACGCGGAAGCTCATTTGATCCGCTTCAACAATAG 1353
 Db 1586 CAACATCTTTAATTTAATCCAAACAGTCCAAATTAAGCCACATCTGCTGTAG 1645
 QY 1354 CTGTAAGTATCATTAATTAATTAATCAATGACACCAAGCTCTCAAGATTAACAGACA 1413
 Db 1646 CTGTTAGTAAATTAATTTGACACACACATCCCTAGTTAATACGTCTAATCAAGACA 1705
 QY 1414 CTGTAATCAATCAAGTCAAGAGCAATTAAGTCCACATCTGTGAGTGTGCTGC 1473
 Db 1706 ---ACTGACTCTTAAGTCTGTGCAATTAAGAGCGCTGTTTGAATTTTGAATTTGC 1762
 QY 1474 TCTTGAAGATGAGCAAGAGCGTTGGAAGAAATGCAACGCTAAGAGACACT 1533
 Db 1763 TTTTAAAGATGAGAGACATTAATCACTAGAACACATATACATTAAGTGAACA 1822
 QY 1534 TAAATAATCAATTTGCGATGAGCTGAGTACAGCGCGCATTTCAAGATAGCTTAT 1593
 Db 1823 TAAAGAGGAATACATATGCTGATTTCCGAAGCAAGAAATGAAGCCAGAAATGTACT 1882
 QY 1594 GGGCTTTAGGCGTCACTTTCCAGAGCTGGCGGATCAATATATGAAACATTAGAATAG 1653
 Db 1883 GGGGTTTCAAGTCACTTCAAGCAGAGAGACACTTGTACACACACTTGTGAGTCT 1942
 QY 1654 CTGCCAGCGCGC 1666
 Db 1943 CCAACAGAAAGC 1955

RESULT 7

US-09-270-767-30401
 Sequence 30401, Application US/09270767
 Patent No. 6703491
 GENERAL INFORMATION:
 APPLICANT: Homburger et al.
 TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
 FILE REFERENCE: File Reference: 7326-094
 CURRENT APPLICATION NUMBER: US/09/270,767
 CURRENT FILING DATE: 1999-03-17
 NUMBER OF SEQ ID NOS: 62517
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 30401

Db 301 AGTAAAGGCGCATGCCCCATCTAATGTAGCAGAAAACTATTATACATTAAGAGGAGAAC 360
Qy 4823 TAAAGCAGTAATTCGCCAAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 4882
Db 361 TAAAGCAGTAATTCGCCAAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 420
Qy 4883 TCACATCAGTTGACGTCTTTTTCACAGCTCTTAATTAACGCTGATTCGTTTAAT 4942
Db 421 TCACATCAGTTGACGTCTTTTTCACAGCTCTTAATTAACGCTGATTCGTTTAAT 480
Qy 4943 TGAAGCCTATATACCGCTCTTTATGAGAACTAGAGCTGCTTTTCTACCTTTGTTTC 5002
Db 481 TGAAGCCTATATACCGCTCTTTATGAGAACTAGAGCTGCTTTTCTACCTTTGTTTC 540
Qy 5003 AGGCGACTTGTATCCCTTATTTCCGCTCACTAATTTCTAATTAATTAATTAAT 5062
Db 541 AGGCGACTTGTATCCCTTATTTCCGCTCACTAATTTCTAATTAATTAATTAAT 600
Qy 5063 AGCATTCGCTCTTACACTGAACCTATTTTGTATTAATTAATTAATTAATTAAT 5122
Db 601 AGCATTCGCTCTTACACTGAACCTATTTTGTATTAATTAATTAATTAATTAAT 660
Qy 5123 AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5145
Db 661 AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 683

RESULT 4
US-09-270-767-21300
Sequence 21300, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 21300
LENGTH: 1420
TYPE: DNA
ORGANISM: *Drosophila melanogaster*
US-09-270-767-21300

Query Match 12.8%; Score 661; DB 3; Length 1420;
Best Local Similarity 99.0%; Pred. No. 3,4e-186;
Matches 676; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
Qy 4465 TGTACATCAGAAAGCAGCGAACTGCATCATGTGGGAGAGAACTCTCAAGAACTCCT 4524
Db 1 TGTACATCAGAAAGCAGCGAACTGCATCATGTGGGAGAGAACTCTCAAGAACTCCT 60
Qy 4525 CCGCGCATCGTGTGATGATGATGCGGAGCCCTTAATAGATTCCTGC--TCGTGCACC 4582
Db 61 CCGCGCATCGTGTGATGATGATGCGGAGCCCTTAATAGATTCCTGCATCAGACACACC 120
Qy 4583 ACAAACAAGACAGACCGGGTTCCTTCCCTGGCTGAGAAAGAGTGAAGAGCGGCG 4642
Db 121 ACAAACAAGACAGACCGGGTTCCTTCCCTGGCTGAGAAAGAGTGAAGAGCGGCG 180
Qy 4643 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4702
Db 181 ACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 240
Qy 4703 CGTACTGTGTACTGCTCGATGATATATTCAGCAGATGCAATGCGGCCGAGAAAGA 4762
Db 241 CGTACTGTGTACTGCTCGATGATATATTCAGCAGATGCAATGCGGCCGAGAAAGA 300
Qy 4763 AGTCAAGGCGCATCGGCCATCTTAATGTGAGCAAGAAATTAATTAATTAATTAATTAAT 4822
Db 301 AGTCAAGGCGCATCGGCCATCTTAATGTGAGCAAGAAATTAATTAATTAATTAATTAAT 360

Qy 4823 TAAAGCAGTAATTCGCCAAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 4882
Db 361 TAAAGCAGTAATTCGCCAAAGTGTACTAAATTTGAGCCAACTCCGCAATCTCACTTC 420
Qy 4883 TCACATCAGTTGACGTCTTTTTCACAGCTCTTAATTAACGCTGATTCGTTTAAT 4942
Db 421 TCACATCAGTTGACGTCTTTTTCACAGCTCTTAATTAACGCTGATTCGTTTAAT 480
Qy 4943 TGAAGCCTATATACCGCTCTTTATGAGAACTAGAGCTGCTTTTCTACCTTTGTTTC 5002
Db 481 TGAAGCCTATATACCGCTCTTTATGAGAACTAGAGCTGCTTTTCTACCTTTGTTTC 540
Qy 5003 AGGCGACTTGTATCCCTTATTTCCGCTCACTAATTTCTAATTAATTAATTAAT 5062
Db 541 AGGCGACTTGTATCCCTTATTTCCGCTCACTAATTTCTAATTAATTAATTAAT 600
Qy 5063 AGCATTCGCTCTTACACTGAACCTATTTTGTATTAATTAATTAATTAATTAAT 5122
Db 601 AGCATTCGCTCTTACACTGAACCTATTTTGTATTAATTAATTAATTAATTAAT 660
Qy 5123 AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5145
Db 661 AAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 683

RESULT 5
US-09-270-767-14266
Sequence 14266, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 14266
LENGTH: 466
TYPE: DNA
ORGANISM: *Drosophila melanogaster*
US-09-270-767-14266

Query Match 8.8%; Score 452.4; DB 3; Length 466;
Best Local Similarity 99.6%; Pred. No. 3,2e-124;
Matches 464; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 2154 GGAATACCCCGAAGCAGCAGCTCTAGGAAAGATCCAACTAGTCAAGTGGTGCC 2213
Db 1 GGAATACCCCGAAGCAGCAGCTCTAGGAAAGATCCAACTAGTCAAGTGGTGCC 60
Qy 2214 TTGATGAACCGATATGATCTTACAGTGGGGGTCTCGACGTACCGCCGAGAGAAC 2273
Db 61 TTGATGAACCGATATGATCTTACAGTGGGGGTCTCGACGTACCGCCGAGAGAAC 120
Qy 2274 AACCAGTAAGCAATCGCGCCGCGACAG--CTGCTGGCGGAATCCCGTAAGCAGAAACA 2332
Db 121 AACCAGTAAGCAATCGCGCCGCGACAGCTCTGCGCGCAATCCCGTAAGCAGAAACA 180
Qy 2333 TACATTAAGCGTGGAGATGATGACAGCAGACTATGTTTCCGGGACCTACATCGCAG 2392
Db 181 TACATTAAGCGTGGAGATGATGACAGCAGACTATGTTTCCGGGACCTACATCGCAG 240
Qy 2393 CCGCGGAATGGGATGGGTAGAGAGCTATAGGAGCTATAGATGTATGACATGACTC 2452
Db 241 CCGCGGAATGGGATGGGTAGAGAGCTATAGGAGCTATAGATGTATGACATGACTC 300
Qy 2453 CGAGCCAGTCTGTGTGTGATGAGAAAGATCTTTGAGCTCAGCTACCTAGAGGTAAATA 2512
Db 301 CGAGCCAGTCTGTGTGTGATGAGAAAGATCTTTGAGCTCAGCTACCTAGAGGTAAATA 360
Qy 2513 ATCAACTACTCACTTAGCGGAGCCACACCCGCTGGAGCTGAGACGACGCGGCGCC 2572

```

QY 579 RSISAVDTAAORAKVRAQYTLYSRQKPLGPNNSNOA-----SMTGAAAGSLPRPL 632
DB 1486 RSLP-----TYS--YKPFSTRPYQSWTAPITVGPAGSGF---TSL 1524
QY 633 NSNSGTPATP-----GSVTPRP---RGRAGVSQSQPGSRSTSPYKLRDQYGGI--- 680
DB 1525 SSSSNTSPASPLKSIWSVSTPSPKSTLIGASTSSVKSISDVASPIRSLRTWSPIKTV 1584
QY 681 -----GNYR--GATGAIKKAGIPRSTASSRETSPTRSGGLMKRSMYSTG 726
DB 1585 VSGSPYNIQVSSGTLARAPAVTEATPLK--GLASNSTFSRSTSPVTTAGSLLERSSIT-- 1640
QY 727 AGSRATPERNNPVRPSPAPALLAQSRAEHTLGVDDGQPDYVSGDYMRSGMRMKLM 786
DB 1641 ----MTP----PASPKNIMMYSSSLPFK-----SITSAAPLISPLK 1676
QY 787 G-----RDESDIDSEASVCSERSPDSYTRGNKSNYSLSGSHRLDWSGORAPPDIE 841
DB 1677 SVVSPYKSRVDYISSAKITWASSLSSPVKQMPGHAELVALNGSISPLKYASS-----S 1729
QY 842 TIIQFCASTHMERKDKGLISLTQYLDGKELTQOOLKCVLDNRKMFMDTHTKVYSLFD 901
DB 1730 TLINGCKAT-----ATLOEKISSATSVSSVSAATDVEKVFSTTTAMPFSPILRS 1780
QY 902 TYTELLVHANETSRNGSSCLTRLFNKLTGDLNMSHKIMKTLQVHEHYPTQOLKE 961
DB 1781 YVS--AAPSPQSLRTPSASA--LYTSIGSS--ISAITSVTSITITVPYVSVNVLPPEP 1834
QY 962 LFRILSDSTQPTTKTRIALRFLTDLA-----NTYCKSSDPFSDQACERTVLKLAOL 1016
DB 1835 ALKKLPDUNS--FTKGAALLSPIKTLTETHQPHFRSTSSPVKSSLFLAPSAKLTSTP 1892
QY 1017 AADQSMELRSQARSC-----LVALYNLTPOMTLLADLP-----KXY 1055
DB 1893 SSISSSOEILKQVAENKEDLMRTAILQTDVPEEKPEOPELPEKGRIDDEEPFKIVEKYK 1952
QY 1056 ODSAR-----SCISHMRQSQSCNGANSPLSSSPKPLQSPSVGFASLOS 1107
DB 1953 EDLVKQSEILKQDVQDNKSGPKSPKSDKG--HSPEDDWIEFSSEETREA----- 2000
QY 1108 HHHQLSISSTSPRSQSVQELLFSSSELDIQNIQKTEBEIRHCFQGOYOTALAPNGFN 1167
DB 2001 --RQQAASQSP--SLPERVQVAKAASEKD--YNLTQVIDYLTNDIGSSSLT----- 2047
QY 1168 GHLOQ--HDQGOQDSCASLSNSKTQSSANTQSTPESATMRLDNLERE----- 1215
DB 2048 -NLKYFEEDAKKQBGQKRVLPALALQEHKLMKMP--PASMRTSTSEKEICKMADSPFGT 2105
QY 1216 -----RTQNAKSPPTDAKIVTS---INMAENGELILASNLMESEVVRVALTLT 1262
DB 2106 DTLESPPDPSQHDQKSPLSDSGFEETRSEKTPSAPQSAETTGPKPLFHEVPIPPVITET 2165
QY 1263 KDQPVLELQTSLTNLGICIKGNCCELPN-----KHFRS 1295
DB 2166 RTEVVAVIRSYDPSAG-----DVPQOPPEEPVSPKPSPTFMELEPKPTSSIKEKYK 2217
QY 1296 IMRMLNIIIEAETHTDVVIAGLVLSKIMSNKRRHMMHFLLELILKIIQCYOHSKEALR 1355
DB 2218 AFQMKRASSSEDDHN-----RVLSTKGMRYKEETH-----ITTTTMYVHSPDGE 2261
QY 1356 DIDSMIPRIAPSLPLDLSTINIVPVATGEFFTNLCAILLEVTGHGSEITDAHLDIY 1415
DB 2262 GASERIEE-----TMSVHDIMKAFQSGRDPF-----KELAGLFHFKSA-----V 2300
QY 1416 FPNLASADDTQGVKRAAVFCIVKLYFVLGEEKVVKPLSVLNPSKVRLLNVYIEK 1471
DB 2301 SPDVHSAASAEKTSQHAHEK-----DNQMKPKLE-----RIIEVHIEK 2336

```

QY 576 PSMRISAVDTAAQAKAVAOYTLVSRKPKPLGPNNSQASMTGAASGLSPRLNSN 635
 DB 622 PSMNPSSSTPGSSQSTTTEGSSASS---PTGSGTSTPSVATEVTSQSTVP---SGS 673
 QY 636 SGGTPTAT---TCGSVTPRPRGAVGQSQP-----GSRF--TSPFKLADQYGGIG 681
 DB 674 SLGTOSTNSPSSPSSLSPTSGMSTLT--SEPSPSSTOSSGAQSTLTTPSPNPSQSTSSLE 732
 QY 682 NYVRGATGA-----IPKASGIPRSTASSNETSPTRSGGGLMKRSMYSTGAGSRRT 732
 DB 733 SSTSGATTSAGSAGTTWTSPSSSVSSQSGSTSPAASTTSG-----EMTSQSTGT 784
 QY 733 PEANNVPRBPAPRLAQAQREAEHTLGVGD--GQPDVSGDVNRSGGMRGRLMGDRDS 791
 DB 785 PGSS-----VSTSAAILTSTQOQSVTNSPGSTVTRPSTVSGSTSSGSTVTVG-----S 832
 QY 792 DDIDSEASVCSRSPSSYTRGNKSNYSLSGSHTR-----LDMGTORAPPDIE 841
 DB 833 TEASTGSSVASSSPADSTQONPVPSTSSGSSMITQSPYSGSTSPVESSTTPSPGSPGT 892
 QY 842 TTIQPCASTHSEKQGLSLTD--YLADGKELTQOQKCYLDLFRKMFMDTHTKVYSLFL 900
 DB 893 TL---TSPSPSSQSTTIGSTGSGTSPGISTSEE-----MTSGSGTQTPGSGTG 938
 QY 901 DYTTELILVHANET-----SRNGSSCLTRLFNKLGTDLNSMHSKIMKTLQYVHXY 952
 DB 939 STVTOPSTVSDSTSGSTVTVGSTEBSSPFIS--TSQNTNPTSTSSGS-----984
 QY 953 FFPQQLKELFRISDSTPPTTKTRIALRFLTDLANTYCKSDPPSDQSQACERTVLK 1012
 DB 985 -----SMSTQTPQSSOSTSPVESSTSGNTS---SGSGPG-----TTLVS 1020
 QY 1013 LQAQADQKSMELRQAQSCVLAALY--NLNTPQMTLLADLPKYQDSNRSCINSHMR 1069
 DB 1021 ISSPSPSSSTIGSSQSTSPVVGTTISQGETETGSGTSTVTKPSTVSGASSSGSTATMG 1080
 QY 1070 QSQSCNSGANSPPSSP-----LSSSPKPLQSPVCG-----P 1101
 DB 1081 TEASSTSGSSTSPSPSQSTSPSTSGATSSPGSSGTTLTSTSPSPQSSSTIGSSQSTSP 1140
 QY 1102 FASLQHHHQLGISSTSPRSROSSVQ-----ELFPSSFLDIQHNQ-----1143
 DB 1141 VVSTTGDMTSGQSTQIPGSGTSTVTPSTGSGSTSGEITSGSTQTPRSSLSTSPAI 1200
 QY 1144 KTSEERHRCFGQYOTALAPNGFNGLQYHDQOQDSCLASSNSKT--OSSANTTOSNTP 1202
 DB 1201 STSTQSVSTNSGSTVTPSTVRSST--SSGSTVTTGSTBSSSTSGSSSATSSSSSP 1257
 QY 1203 ESATMRDLNERERTTONAKSPDDAKVITVSINMANGEL 1243
 DB 1258 VSTSGSPN---PSTSGSTPTPNPSQSTSPVSTTTGEM 1294
 RESULT 15
 A55575
 Ankyrin 3, long splice form - human
 N:Alternate names: ankyrin G
 C:Species: Homo sapiens (man)
 C:Date: 03-Mar-1995 #sequence, revision 03-Mar-1995 #text_change 31-Dec-2004
 C:Accession: A55575
 R:Kordeli, E.; Lambert, S.; Bennett, V.
 J. Biol. Chem. 270, 2352-2359, 1995
 A:Title: Ankyrin-G. A new ankyrin gene with neural-specific isoforms localized at the ax
 A:Reference number: A55575; PMID:5138209; PMID:7636469
 A:Accession: A55575
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-4377 <KOR>
 A:Cross-references: UNIPROT:Q12955; UNIPARC:UP1000013C497; GB:U13616; NID:g608024; PIDN:
 C:Genetic: GDB:ANK3
 A:Gene: GDB:ANK3
 A:Cross-references: GDB:424503; OMIM:600465
 A:Map position: 10q21-10q21

C:Superfamily: EGF homology
 C:Keywords: alternative splicing; peripheral membrane protein
 F:73-105/Domain: ankyrin repeat homology <AN01>
 F:106-138/Domain: ankyrin repeat homology <AN02>
 F:139-171/Domain: ankyrin repeat homology <AN03>
 F:172-200/Domain: ankyrin repeat homology <AN04>
 F:201-233/Domain: ankyrin repeat homology <AN05>
 F:234-266/Domain: ankyrin repeat homology <AN06>
 F:267-299/Domain: ankyrin repeat homology <AN07>
 F:300-332/Domain: ankyrin repeat homology <AN08>
 F:333-365/Domain: ankyrin repeat homology <AN09>
 F:366-398/Domain: ankyrin repeat homology <AN10>
 F:399-431/Domain: ankyrin repeat homology <AN11>
 F:432-464/Domain: ankyrin repeat homology <AN12>
 F:465-497/Domain: ankyrin repeat homology <AN13>
 F:498-530/Domain: ankyrin repeat homology <AN14>
 F:531-563/Domain: ankyrin repeat homology <AN15>
 F:564-596/Domain: ankyrin repeat homology <AN16>
 F:597-629/Domain: ankyrin repeat homology <AN17>
 F:630-662/Domain: ankyrin repeat homology <AN18>
 F:663-695/Domain: ankyrin repeat homology <AN19>
 F:696-728/Domain: ankyrin repeat homology <AN20>
 F:729-761/Domain: ankyrin repeat homology <AN21>
 F:762-794/Domain: ankyrin repeat homology <AN22>
 F:795-827/Domain: ankyrin repeat homology <AN23>
 Query Match 2.5%; Score 192; DB 2; Length 4377;
 Best Local Similarity 18.1%; Pred. No. 0.11;
 Matches 303; Conservative 242; Mismatches 613; Indels 518; Gaps 76;
 QY 69 SLEAFSELIKRGSPFNATATVLPVHVIDRLGDSRDTREKQOLLRLDMERVLPRQL 128
 DB 906 SLGARSASLRSFSS--RST-----LNRSYADSMWTE--ELVPSKEQLLTFPREP 956
 QY 129 IDKL-----ATSCFGKHN---AKVREEFLOT--IVNAL--HEXGTOOLSRYVYIP----- 172
 DB 957 SDSLHYSWAADLTLDNVLVSPHSGFLVSMVDARGSMGSHHGRIIIPPRKCTA 1016
 QY 173 ---VCALL-----GDPTVNRERAIQTLVEIYHVGDRLEPDLRRMDVP----- 214
 DB 1017 PRTICRLVKKRKLANPPEHGRGSISSLVEMGPAQFLQPLV---VEIHPGSMRCK 1073
 QY 215 ASKAML-----ECKPDVQKQEGLLPRLAKNTNGNGVGLDEADNIGLRERTRM 265
 DB 1074 ERELVLSENGETWKEQFDS--KNEDL-----TLNLMDELDSPBELG--KKRICRI 1126
 QY 266 KR--PLHSAVSSSLRKPKN--VNDVTGDAQAVTMESFESEF-----EVVPO 307
 DB 1127 TQDFQYRAVYSRIQESNOIQPEGGILSSTTVPLVQASFPGALTKIRVGLQAPV-- 1185
 QY 308 LNIFFAKMDIYKQVL-----V1SDKNADWEKRV----- 338
 DB 1186 -----DEIVKILGNKATFSPITVTEPRRRKFKITWTIPVPPSGSGVNGYKA 1236
 QY 339 DALKKIRALLISYTPQPOFAVQKELS---LSFV-----DLT 374
 DB 1237 DTPNMLRLCSITGQTSPP---AQWEDITGTLPLFFIDQCSPTTNVSARFMLADCHQVL 1292
 QY 375 KE-ELRSQVIREACTTIAYMSK-----TLRKLDLAFCS--ILEHLINLIQNSAKY 422
 DB 1293 ETVGLATQYRE-LICVPYMAFVVPAKNDVSSLRFCCTDDKVDKDTLQOQENFEVY 1351
 QY 423 IASASTIALK---YIKYTAAPKLLK-----IYDTLQO---SKSKDIRSTLCEL 466
 DB 1352 ASKQIEVLEKPIVDCGNLAPLTKGQQLVFNFYSFKENRLPFSIKIRDTSGPCOR 1411
 QY 467 MYLFEENQYKLENNATV--LQDTL-----KSIQADCDARHSHR---YAYVAFRRHPE 518
 DB 1412 LSFLEKRTTKQLPQVACNNLITLPAHKKETESQDDDEIKTDRQSFASLALRRRSY 1471
 QY 519 LADQVGTLDIAQAALREBEREGGGGCGTGTGAPETRRIVSRIGRTGLQKPTPM 578
 DB 1472 LTE-----PGMIERSTGAT 1485

QY 1076 SGANSFSSPLSSSPKPLQSPVGPASLQSHH-----QUSIS-----STSPRSQSSVE 1127
 DB 4273 SRPESVASQP--ESVSPSPSGAA-----SHEHKEVEISESHKAEKSRPESVASQVS 4322
 QY 1128 QELL-----FSELDIQHNIQCTSEIRHCFCGOYOTALAPNGFNHLOYHOGGQDS 1180
 DB 4323 EKDKTRRSPASSTQSFSTKEGDEETSLH-----SLTTTEVETKOMEKESPFES 4374
 QY 1181 CASLSNSKTQSSANTQSNTPESATWRLDLERERTONAKSP---TDDAKVTVSIM 1237
 DB 4375 VSTSVTSSTVLSQSTQ-----LREESTSESLSSLKXEDSRRRSLSL 4421
 QY 1238 AENGEELLASNLME-----SEVVRVALTTKQDPVELLOTSLNLGICIKGNCEL 1288
 DB 4422 AEKGIGATNTSLKEDTASASQLEELVQSECESESIVSEIQTSI-----AQK 4470
 QY 1289 PNKFRSRIMMLNLILAEHTDVYIAGLHVLSKIMSRKMMHMLLELILKICQY 1348
 DB 4471 SNK-----EIKARERTKTSQFTTTTSATKQDSLKEVIAEF--LATEKIVSAKE 4518
 QY 1349 -HSKEALRDIIDSMIRIAPSLPLDLISINIVNPVATGEFPTNLCAIKILLEVEHSGSEI 1407
 DB 4519 AFTSEATKRSADCLKTTTASVSTSSASQRLPFGTDBSRRE---SLSQASB---SRL 4571
 QY 1408 TDAHLDIVFPPLASADDT-----OSMVRKAAVFCIVKLYPVLGEEKVYKPKLSVINPS 1460
 DB 4572 T--HSD---PEDEPADVDVDRSSVSKESRSKIATIMTSTY-----KPSDEMPIS 4618
 QY 1461 K-VRLNLVYIEKQNCISGGGSSST-----KNSSAASS 1492
 DB 4619 KLVEEHEHVEELAEQVTSKTTLLQSSSEOSTTSS 4657

RESULT 11

Probable transposon protein (imported) - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear crease)
 C/Date: 02-Mar-2001 #sequence_rev10 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: F86403
 R: Theologos, A.; Ecker, J.R.; Palm, C.J.; Federpohl, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.M.; Chung, M.K.; Com, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Hutzar, L.
 Nature 408, 816-820, 2000
 A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A: Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A: Reference number: A86141; MUID:21016719; PMID:11130712
 A: Accession: F86403
 A: Status: preliminary
 A: Molecule type: DNA
 A: Residues: 1-1148 <STO>
 A: Cross-references: UNIPROT:Q9G6N3; UNIPARC:UPI000009EBB8; GB:AE005172; NID:g10998923; F C: Geneticks:
 A: Map position: 1

Query Match 2.6%; Score 198; DB 2; Length 1148;
 Best Local Similarity 20.0%; Pred. No. 0.0061;
 Matches 233; Conservative 136; Mismatches 432; Indels 362; Gaps 56;
 QY 516 FPELADQ-----IYGTLDIAQALEREHEGGGCGTGTGTAPET----- 557
 DB 108 FPLSDDPAPASVVRGRPOSQISLSRSSTWKSRSKSGASPNRLSTSRANMOQIR 167
 QY 558 -----VSRIGRTPG-----TLQKTPSMESISAVDTAAQRAKVR 595
 DB 168 GRPSAHPSPASGRSGTTPVRRISPTGKSPGSPVRSPTLSRMSTGSTMSPA--- 224
 QY 596 AQTLYSRQRKPLGPNNSQASMTGAASGLP-----PRLANSNGGTATTPGAVTP 649
 DB 225 -----VAGTSPVSSSRGNSPSPKIKVQSNIPGSLDAPNRLTSLDRPASVYRGSSP 278

QY 650 RPR-GRAGYSQ-----SQPGNSRSTPS--TKLRDQY-----GGICNY-----Y 684
 DB 279 ASRNGRDAVSTRKRSVSPASRSVSSHSHEDRPSQSKGASVSGDDLHSLQSI PV 338
 QY 685 RGATGAIIPKAGSLP--RSTASGRSTPSRGGGLKMRMYSTGAGSRRTPERNN-----P 738
 DB 339 GGSERAVSKRSLSPSRSTSRSSKLSLSP-----GSAPRRPFEALQMEHPKHHMFRP 393
 QY 739 VRPSAPARLQAQREA-----EHTLQVGDGQPDVYSGDYMSSGGRMRKMLGRD 789
 DB 394 LASSLSTGTISGKSSSYHHIMLRHSTATVGSNSSGQYTGMPDAKGM-----D 444
 QY 790 ESDDIDSEASVCSSESFPSSYTRG--NKSNTLSGSHRLDMSTQAPRFDIETITQFC 847
 DB 445 PVPFQSEVENLAVPDKHEESIAFGVNVNLSNESHSHESPSDQIGMDODYTV--EC 502
 QY 848 ASTHMERKDGSLSLQVYADGKELTQOOLKCYLDHFRKMFMDTHFKVY--SLPLDTYTE 905
 DB 503 ES-----SANEVSHQ-----VPDENSS--STHSLHVGNEFLBQV-- 536
 QY 906 LILVHANET-----SRNGSSCLTRLPNKLGTDLNLSHMKIWTLOVHEVFPQOLKE 961
 DB 537 -----ALETMEVCGRGSHYCATE-----ATRSEINICECREHSHFVEDT----- 577
 QY 962 LFRITDSTQTPPTKTRIALRLFLTLANTYCKSSDPSPSQSQCERTVAKLQADLADQ 1021
 DB 578 -----SPGTNSP-----KLSQTFDEN 594
 QY 1022 SMELRQARSCLVALYNLNPOMTLLADLPKY-----YDSARSC 1062
 DB 595 K-----LTFENLVID--VLDSLPVYVVEELIETBEKIEQCNSSYEQEYHL 640
 QY 1063 IHSMEROSQSCNSGANSPPSSSPKPLQSPVGPASLQSHHQLSISTSPRSR 1122
 DB 641 YESISIRALEQOVDMNLNYDQSSGCGPLISIGTKDTQGLDKHDVNVISLGRGDV 700
 QY 1123 QSSVEQELLFSSSLDIQHNIOKTSEIRHCFCGOYOTALAPNGFNHLOYHOGGQDS 1182
 DB 701 PLVTKSVSKSPV--IOANNS-----CFTRSYE-----GFS-----YSRDRSI 738
 QY 1183 SLSNSKTQSSANTQ-----SMTPEATWRLDLERERTONAKSPFDDAKVTVS 1234
 DB 739 SLRSSTELTASASSMDYGSIRKGSHTRQSSGTLD--LEHRTVDTNKS-----LSTMS 791
 QY 1235 INNAENGELLASNLMS--EVRVALTLTKD-----QVELLOTSLNLGICI 1281
 DB 792 SSSGMSHTQALNWPEDSFEMCAQMTCTLDBTHQESHTEPQNLCKETNVNADFE 851
 QY 1282 KGCNCELPNKHFRSRIMMLNILE--AEHTDVIY-----AGLHVLSKIMSRKMMHML 1333
 DB 852 SVG-----LVRIANVLGLDLAEHNPVYVMDDECCENGDNVANTVLSKGTRES 900
 QY 1334 HFELILIKI-----QC--YQSHKALRDIIDSMIRIAPSLPLDLISINIVNPV-----AT 1383
 DB 901 HIRSTDGLASPTTDDCPFNDSRLQENDVNETPHGLSTTASIEPESSEPELPGICVH 960
 QY 1384 GEFF-----TNL-----CAIKILLEVT--EHHSS-----EITDAHLDIV--FPN-----LAR 1421
 DB 961 DELPESRRLNANVDDGSEKMYASVDHSSSAVNEILDESITLVPCPGKEPRSLTLE 1020
 QY 1422 SADDT-----OSMVRKAAVFCIVKLYVLGBEK--VKRKLSTVNLSPKVRNLNVYIEKQ 1472
 DB 1021 EATDTILFCGSIYHDLVYQAATITAMKAKDVPAEBEMLHPTVVLGKSN-----N 1071
 QY 1473 RNCIS--CGGSSYK--NSSAASS 1492
 DB 1072 RNSYGLGGGTAKKRSKAAKAS 1094

RESULT 12

C84586
 hypothetical protein At2g20200 (imported) - Arabidopsis thaliana


```

QY 747 LLAOSREAEHTLGVDDGQPDVYSGDYMRSGMKRMKMGKRGDESDIDSSAVSVCERS 806
DB 611 LKNKIQKYE-----ILSG-----SSVDLPSEAFELSSNL 640
QY 807 FDSYTRGNKSNYSLSGSHTRLDWSTORAPFDIETII-OFCASTHMERDGLISLTQY 865
DB 641 TDALYS-GSSICYSILFISHLLDLTFQ---YVDIASLSQPLLCVYDPSNVGSHFALASF 696
QY 866 -LAQKELTQOQLKCVLDMFRKM- MDTHYKYSLFDLTVELILVHNETSRNGSSSCL 923
DB 697 PYVSHYDAHKYFPPIVEFVLNINISMAHVKVP--FVTHQKRIIH-----GCL 744
QY 924 TRLNKLGTDLNSHMSKIKTKTLOVHVEYFPQOLKELFRITISDSTQPTTKRIILR 983
DB 745 LML-KEISDTKLANDLENK-----PFVYT-----DKLRYSSKILAMAKTKLP--- 786
QY 984 FLTDLANTYCKSPDPDSQACERTVLAQLAADQKSMELRQARSCLVALYN-LNTP 1042
DB 787 -----SKWNIPLSGLL-----FSLRAHDTFMFDGLDLRLNEE 818
QY 1043 QMTLLADLPKYVDS---ARSCIH-SHMRQSCNSGANSPESS-----PLSSSSPK- 1092
DB 819 SRTYLVASMSK--QDAPDYKSSTHQEHLSKNLPTLNTSSSNSQTDLLVPHCKGETKE 876
QY 1093 -PLQSPVGPASLSQHHLQSLSTIS-PRSSQSVQELFESSLDIQHNIQKTS 1146
DB 877 TENQSPLESKEGLSKDTHIESPGSTLEKENEENEGKNPYESNC--SEESLDHNDIQTL 934
QY 1147 EELRHCGQGYOTALAPNGFNHQLQYHDQGDSCASLSNSKTSQASANTQNTESA- 1205
DB 935 VNKKETLAQDSESLQGNALNEKGFENQGLSSAAKYLKCDTLQHVSPISNSVSSP 994
QY 1206 -----TWRLDNLERERTTQ---NAKSPTDKAYTVSINNAENGCELILASNLMESEY 1254
DB 995 KDFTRTPFEINGERETGFELSYVALSKKD---INVQKTEVDESVEGIANAFMDN 1049
QY 1255 VARALLTLTDQPV---ELQTSLTNGICIKGNCGLPNKHFPSINRMLNITIEAHTDV 1311
DB 1050 VNQDSINSYDQSSGKDKLLTSTST-----PNKPTTFMPANEILLGSPAKDY 1096
QY 1312 VIAGLHVLKSKMNSKMRHN 1331
DB 1097 DHDQSYVHEHLSHMRHN 1116

RESULT 10
113564
microbubble-associated protein homolog - fruit fly (Drosophila melanogaster)
N:Alternate names: hypothetical protein EG:49E4.1
C:Species: Drosophila melanogaster
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 31-Dec-2004
C:Accession: T113564
R:Spanos, L.; Papadimitrakis, G.; Siden-Kiamos, I.; Louis, C.
submitted to the EMBL Data Library, April 1999
A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.
A:Reference number: Z17689
A:Accession: T113564
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5327 <SPA>
A:Cross-references: UNIPROT:O76891; UNIPARC:UPI0000110149; EMBL:AL031128; PIDN:CAA20006.
A:Cross-references: FlyBase:FBgn0025392
A:Introns: 24/2; 52/3; 104/3; 179/1; 232/1; 1669/3; 2566/1; 4798/3; 5272/1
A>Note: EG:49E4.1

```

```

Query Match 2 6%; Score 198.5; DB 2; Length 5327;
Beat Local Similarity 19.3%; Pred. No. 0.064;
Matches 308; Conservative 227; Mismatches 635; Indels 429; Gaps 70;

```

```

QY 102 SRDVTREKAQQLLRPLMLHEHRLVLPQALIDKLATSCFKKKNKAVREFFIQTIVNLHNGRT 161
DB 3280 SRDVAEKSPLASKASR-----PASVASVQDDEAKESKESRRRESVASKSPPLAYKAS- 3333

```

```

QY 162 QQLSVRYIIPVICALGDPFTVNVREAAIQTLVEIYKHVGDRPLRMDVPASKLAML 221
DB 3334 -----RPASV---ASIDDEAKESKESRRESVASKSPPLASK 3367
QY 222 EQKFDQVKGEGILLPSALKNTNGVGLDEADNIGLRERPTR---MIKPLHSAVSSSLR 278
DB 3368 E-----ASRPTVASVQDEAK--SKESRSRQSVAKESPLASKAS--R 3408
QY 279 PKRNVDYGDAGAVMSEFESSFVYVPLNTHAAD-----MDIYKQVLVI 327
DB 3409 PASVASVQDEA---EKSKESRRESVASKSPPLASKASRPASVASVQD----- 3456
QY 328 SDGNADMERKVDALKKIRALLILSYHTQPOFVAVQKELSLSPVDLTKELRQVIREAC 387
DB 3457 AEKSKESRRESVASK--SPLASKASRPASVASVQDEA---EKSKESRRESVAS-- 3508
QY 388 ITAYVSKTLRNKLDACWLSLEHLINLIQNS-----AKVIASASTIALKYTIKYTHAPK 442
DB 3509 -----KSPLPKSKASRPASVASVQDEAKESKESRRESVASKSLASKASRPASVAE 3562
QY 443 LKITYDTLNQSKSDIRSTLCELAVLFEEMQYALERNATVYLDTLKKSIGDADCDAR 502
DB 3563 SVK---DEAKESKESRRESVASKSPPL-----ASKASRPASV---AESVDEAKESK 3609
QY 503 RSRVAVYAFRRHNP-----ELADQYGLTDLAQAALEREERG--GG 544
DB 3610 EYSRRESVASKSPPLSKASRPASVASVQDEADSKSKE---ESRBSGAEKSPPLASMEAS 3666
QY 545 GGTGTGTGTAPEPTRRTVRIQRTPTLQKPTPSMRSISAVDTAAQAKVRAQYTLYSRQ 604
DB 3667 RPTVASVQDEAKESKESRRESVTEKSPLSKSK--ASRPTVASVQDEAKESKEER 3724
QY 605 RK-----PLGNMNSQ-ASM-----TGAASGSLPPRLNSN 635
DB 3725 RESVAEKSPPLASKESRPASVASIKDEAGTKQSRRESPESSGKASIKGDOSLASK 3784
QY 636 SGGTPTATTPGSY---TPRPRGAGVSQSGRSTRSPSTKLRDQYGGIGNYRGATGAI 692
DB 3785 ETSRPSDVSVEYKDETEKPEBSA-IDKQVARSRPASVASAKDEKSPPLHSPRESVADKSP 3843
QY 693 KKAAGIPTS-----TASSRETSPTRSGGGLMKRSMYSTGASRRTPERNP----- 738
DB 3844 DASKEASRSLVAETASSPIEEGPRSIADL-SLPLNTLGEAKGKPLTSSPFDVAEGPFL 3902
QY 739 -VRPSAPARALLAQSREAEHTLGVDDGQPDVYSGDYMRSGMKRMKMGKRGDESDIDISE 797
DB 3903 EVKASSPRPAVLISKPAPFS-----QPD-----TGHYASTPVD-E 3936
QY 798 ASSVCSERSFDSYTRGNKSNYSLSG-----SHTRLDWSTORAPFDIETI--- 843
DB 3937 ASPVLEIEVEQHT--TSQVGATGATAETDLDDLLETYSKSYTVKQSEETLFFETLSKV 3993
QY 844 -----IOFCASTHMERDGLISLTQYLDGKELTQOQLKCVLD----- 882
DB 3994 ESKVEVLESSVQYQEVKQTSVQAETTVDSLQQLTK--KSSQGLT--EIKSVLDNTIS 4049
QY 883 -----MFKMMMDHTIKYKSLFLDTVITLILVH-----ANETSRNGSSSCLT 924
DB 4050 NVTNLFSTAVETIEKKVQDVYTERK-----IEKATHEVSEHVTYTGESSSTETSQEKSSLDL- 4104
QY 925 RLPLKLGTDLNSHMSKIKTKTLOVHVEYFPQOLKSE-----LFPILISST 970
DB 4105 GTFELMETHHTYGSSEF--TVITCEDEPVLHDIKEDEBHRSPSPSDVUKAIIIPPQP 4163
QY 971 QPTTKRIAILRFLTDLANTYCKSDPS-----DOSQACER-----TVLKLQAALAD 1019
DB 4164 MRPLSPREEVAVAKIVADAVAKLSKDQITDIIPDFERQLEKUKSTADFEEDSKTRD 4223
QY 1020 QKSWEILSQARSCIVLYALYN-LNTPQW---TLILADLPKYVQDSARSCISHMRQSCSN 1075
DB 4224 EKSLDI-----SVKVEIESKSSPDQKSGPISIEKDKIBQSEKAQUL-----ROGILAS 4272

```

RESULT 8

S28293

hypothetical protein ZC94.3 - *Caenorhabditis elegans*C/Species: *Caenorhabditis elegans*

C/Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004

C/Accession: S28293

R/Thomas, K.

submitted to the EMBL Data Library, December 1992

A/Reference number: S28285

A/Accession: S28293

A/Molecule type: DNA

A/Residues: 1-643 <TPO>

A/Cross-references: UNIPROT:Q03609; UNIPARC:UPI000017BCEA; EMBL:Z19157

C/Genetics:

A/Intons: 141/2; 167/3; 245/3; 341/2; 412/3; 600/3; 640/3

Query Match

Best Local Similarity 19.5%; Pred. No. 5.8e-07;

Matches 140; Conservative 137; Mismatches 265; Indels 175; Gaps 28;

```

QY 796 SEASSVSESRFSDSYR-----GKSNVSLGSHRLD-----829
DB 40 TRTSITSDSRDTSPTRRNSPLPETOKARVKYNGSFFAKLGMPDPTDDEFLPIRI 99
QY 830 WSTORAP----FDDIETIIQFCASTHMSERKGLISHTOYLADGKELTQOOLKCVLDMFR 885
DB 100 RSLPKTTIEAHDKVQVLEKCCSSVSKEKGIKLLPIYAD-TSLNPIEIKNIGCLN 158
QY 886 KMFMDTHT---KYSLFLDTVELLIVANETSRNGSSCGTLTFENKLGTDLLNSHAK 941
DB 159 RLISDAANTWLEIYSIFVFT-----HSSRLS-EWLRLLALXLFARKAETLPNTKQ 210
QY 942 IWKTLGVHVEFFPQOLKELFRISDSTOPTTKTIALIRPLTDLANTYCKSSDPSD 1001
DB 211 IGHLLNLTLECFNAHDLVTCELMCDPIHLMVPKARVLELYTSLIDEYTEPG-AS1 268
QY 1002 QSQACERTVLKLAQLADOK-SMELRSQARSCLVALYNLTPQMTLLADLPKYVODSAR 1060
DB 269 NAKELTKAIRKMLTWADDPRLSILLTPHVEKALICSMCVAVADFSALISD---DSEOK 324
QY 1061 SCISHNRROSQSCNAGNSPSSPLSSSPKPLQSPVGFPSLQSHHQLSISSTG- 1118
DB 325 NMIHQTQIR-----NGLENGISNNIATNS-----GATASRETSNTSFOKESTFG 370
QY 1119 -PR--SQSSVVEQELLSSSELDIQNTQKTESEIRHCFGGGYOTALAPNGNGLQYHDQ 1175
DB 371 LPEFGARKGTGVNL-----GSLINISNNLALSRLE-----400
QY 1176 GOODSCASLSNSKTSQSSANTTOSNTPEASATMRLDNERERTTONAKSPTDKAVITVS1 1235
DB 401 --EOSTRLMEKXVNLNTV-TLPDPTLEKIONVODLQKRRSENDAE--QESALISS1YM 455
QY 1236 NMAENG-----ELILASNLMESEVVALTLYTDQPEVLLQTSLTNLGICIKGANC 1286
DB 456 MICDGGGVMEQCYAKILL-NLFE-----ILSKSR-----SENNKMC-----492
QY 1287 ELPKHRSIRMRMLNLLEAHTDVVAGLHVLSKIRSNKMRHNMHFLLELLILKLIQC 1346
DB 493 -----LRIGKRCCTAO-----AAKLDPST-----EMAVCKYVDA 521
QY 1347 YHSEKALR--DIDSMIPRIAPSLPLDLISINIVNPVIATGEPFN-----LCAIKILLEVT 1400
DB 522 AVNTNDATLALAVEDCLRTLATHPLSITINIAYIL--NOEPIDDERASIVLKVTRLF 579
QY 1401 EHHGSEITDAHLDIVFPNLASADDTOSMKAAVFCIKUYFVLGEEKVKKLSVL 1457
DB 580 EELPAEELNINIVDDITFTIIKAYOSTSSTVAKTVVCLVAMVNRVGEQRMTPHFTKL 636

```

RESULT 9
T11648probable mitotic spindle protein - fission yeast (*Schizosaccharomyces pombe*)C/Species: *Schizosaccharomyces pombe*

C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C/Accession: T11648

R/Mood, V.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, December 1997

A/Reference number: Z17304

A/Accession: T11648

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1462 <MOO>

A/Cross-references: UNIPROT:Q42874; UNIPARC:UPI000069DB2; EMBL:AL021046; NID:e12167e

C/Genetics:

A/Map position: 1L

A/Intons: 17/1; 39/3

A/Note: SPAC39.12

Query Match

Best Local Similarity 19.4%; Pred. No. 3.1e-06;

Matches 272; Conservative 211; Mismatches 564; Indels 353; Gaps 57;

```

QY 1 MAYRKPSDLDFIQMPKADMRVKVQIADLVTELSDDTNSIVCTDMGFLIDGLMPVLTG 60
DB 1 MADKXADFLKFLKSNASTD--EKTRCLPTLRSEF--NKNININADLGLFVECFRLALTT 56
QY 61 SHFKIAOKSLFAFSELKRLGSDFNAYTATVPH-----VIDRQDSRDVIREKAQLLRD 116
DB 57 VNELLRSSVACFETLRRLRAQYPTWLFKRVYMLKNVLVDHTA-SRDLOKRVNLITL 114
QY 117 LMEHRYLPPQALIDKLATSCFKHNAKVRSEEFQTIY---NALHEGTQOLSRYVYIPV 173
DB 115 LMFNPSEIEKSLIHIST---SKSAETRIQCKRMVLYANALSTDVKSRLALYIN-- 169
QY 174 CALLGPTVNVREAAIQTVEIYKAVGDRLPDLRRMDVPASKLAMEKQPOVKQEG 233
DB 170 ---LENANSVREAEKVELLIYKMLSTSAK--MQFTIVDETTSGLRREILQSLVEBELSL 224
QY 234 LPLSA---LKNTNGCVGLDEADNIGLERPFRMRKPLHSAVSSSLRKPKNVNDVTGD 289
DB 225 ISSSEVIIVQNS-----ASSFQAPMPPTAVATL 253
QY 290 AGAVTSEFSEFVYVPOINIFPAKMDIYKOVLIYSDKNAD--WEKVDALKIRAL 347
DB 254 YQVEVELEN-----VYPLANF-SKQLEODSASMLAFEBERETQKMSVQSDVLRURQY 306
QY 348 L-----ILSYHTQPOFVAVOLKELSLSFVDILKEELSQVIREACITAYNSKTLRNKLA 403
DB 307 LRGNACIDY--LPELLSV-LKTL-LPGILLALSLRTLTSSSAIQLIKEMAIILKSNIDP 362
QY 404 FCMSTLEHLINLIQNSAKVIASASTA-----LKTIITYTAPKLLKIYTD 449
DB 363 FLELILPILNLIKVSIVRKLASQANVTFAAILVNCVLSHNLSEFISLAH-----D 413
QY 450 TINOSSKOIRSLCELMVTLPEEMQTKA-LEENATVLRDTLKKISIGDADCDARHGRYA 508
DB 414 T--NAQLRVSSWMIFNLISLSPELKXLASLQNLKAFELIGGLADSSQVREYVRKS 471
QY 509 YNAFRHFPFLADQIYGLDIAAQRALEREREGGGGGGTGTGTAPETRRTVSRIGRTP 568
DB 472 FMTLSYFPFVQSELNLTLEPSVLKQHL-----499
QY 569 GTLOKPTPSMRSISAVDTAAQAQAKYAOYTL--YSRQRPPLPPNNSNOASMTGAASGS 626
DB 500 --LANPN--RQAASFNFGPKAPRPLSNLNSFSGSQEETSSNSNSSGT-----547
QY 627 LPRPLNNSGCPATTPGVSPTPRPRGRAGVSOSSQSGRSRTSPSTKLADQYGGIGNYRG 686
DB 548 ---RRLGLFORATPABRE-RVLPYTRSQAHSLSLPSLPBSGSPS1-----590
QY 687 ATGAIPKKAAGIPRSTASRETSPTSGGGLMKRSMYSTGAGARRTPERNNPVPSAPAR 746
DB 591 ---AIPS-----KRSVATIKDSKTFE-----610

```

```

Db      1155 --DKSGPIFLVNGEQRNKKDK-GLKYLKNNFTTRBEYIEQLKTKQMSVCVAKWLDE 1211
Qy      806 SFSSTTRGNKSNYSLSGSHTRLDWSTORAPEDDIETIIQFCASHTWSEKDKGLISLTQY 865
Db      1212 MFHSDQOHNNKA-----LAWVD-----HLESEKGVIGC----- 1241
Qy      866 LADGKELTQQOLKCYLDMFRKQMFMDHTVYSLFDVTVELLVIANE---TSRNGSSC 922
Db      1242 -----LDLTKMLTLRFEDNTISVLMKALEYLKLTPTLISEEYHLENENASSF 1230
Qy      923 LTELFNKLG--TDLNSMSHKIKTLQVWHEXFYPTOLKELFRILISSTOPPTKRIA 980
Db      1291 IPLVYKVEBPQVINKVDRAILNRCGLY-----PASKFPPTMGTSKSKSKQRAE 1343
Qy      981 LRFELTDLANTYCKSSDPPSDSQACERT--VLKLAOLADQKSMELRSQARSLVALY 1037
Db      1344 CLEELGLCTVESY-----GMVQCQPTPGKALKEIAVHGIDRRNAVARNALMTITVY 1394
Qy      1038 NINTPQMTLLADLPKVYODSARSCSHSMRQOSQCSGANSPPSSPLSSSPKPIOSP 1097
Db      1395 NVHGQVFLKLNLT-----SEKMSMLEERIK-----SAKRPAAPIKOVEKPORAQ 1443
Qy      1098 SV-----GPFASLOSHHHQLSISSTPSPRSQSVGEOLTPSESLDIQNTQKT--- 1145
Db      1444 NISSNANMLRKQPAEDMSKLNQASMSGHPRAQ-MVRR--FOLDDEIENDNGTVRC 1500
Qy      1146 --SEIRHCFGGQYOTALAPNGFNGLQYH-DQGGODSCASLSNSKTQSSANTQSNTP 1202
Db      1501 EMPELYQHLDDIFEPVILPEPKIAYSPHFDMHSNTASTINFIISGVASGDLNTSIQA 1560
Qy      1203 ESATMRDLNLEERTQNAKSTDDAKYITYSI---NMAENGELLIASMSEVYRV 1257
Db      1561 LTQLFQIESLAREASTGVKQLMHG--LITLMLDSRIEDDEEGQVIRSVLLVVKYLE- 1617
Qy      1258 ALTLFKDQD-----VELLOTSLNTMGICINGGCELPNKRFSRIMNMLNILEAETDV 1311
Db      1618 ----KSDQTNILSALVLLQDLSL-LATASSPKSELMVKCLMWRVRLDPTINSINLDR 1671
Qy      1312 VIAGHLVLSKIMRSNMKNHMHGFLLELLIKIIOCYQHSKEALRDIDSMI-----PR 1363
Db      1672 ILDDIHIMKVPFKPEKTK-----QC--KSEPIRITLTKLHTLCKLKGPK 1714
Qy      1364 IAPSLPDLISINIVPVATAGEFPNLCATKILLEVTEHHGSEITDAAHDIIVPNLARSA 1423
Db      1715 ILDHLTMDNKN-----ESELENAHLCRM-----KISMDOGSKSD---KETAKGA 1757
Qy      1424 D--PTQSWAKAAVCEIVKLYFVLG-EKKVAPKTSVLNPSKVRLLNANYIE 1470
Db      1758 SRIDAKSKAKAVNDF-LAEIRFKIGSKENTKEGLAELEYEKKYSDADIE 1806

RESULT 7
S28294
hypochemical protein ZC84.3a - Caenorhabditis elegans
Cispecies: Caenorhabditis elegans
Cidate: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004
Cidaccession: S28294
RidThomas, K.
submitted to the EMBL Data Library, December 1992
AidReference number: S28285
AidAccession: S28294
AidMolecule type: DNA
AidResidues: 1-730 <THD>
AidCross-references: UNIPROT:O03609, UNIPARC:UPI000017A06B, EMBL:Z19157
CidGenetics:
AidIntrons: 77/3; 120/3; 213/1; 228/2; 254/3; 332/3; 428/2; 499/3; 687/3; 727/3

Query Match      3.6%; Score 277.5; DB 2; Length 730;
Beat Local Similarity 16.0%; Pred. No. 1e-07;
Matches 180; Conservativity 179; Mismatches 334; Indels 429; Gaps 31;

361 VOLKELSLSPVDILKELBSQVIRACITATYMSKTLRNKLDAPCWSILHLLINTLQNSA 420
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

D	b	6	VOU-----LGFELAVNDLSDQVLRBAIAICSFVSKXGJETHSHIGEDILVPMASQAVST	61
Q	y	421	KVIAASATLAKTIITYTHAPKLLKYTTDLNOSKSDIRSTLCMLVLLFEEMQYALE	480
D	b	62	KIMATASATLEPIEVVOTQVFTLLS-SFSTSKDSORROIAALLEIYISKSDRIKK	120
Q	y	481	RNATVLRDTLKSIGADCDARSHSYAAWAFRRHPPELADQIYGLTDLIAQALBRERE	540
D	b	121	QIMROJCELLKSIKINADSETPAAGRAAKADENHSEADALYLELDISKQMLR----	176
Q	y	541	GGGGGCTGTGYTAPETRTVSIRIGTPGTLOKPPPSMBSISAVDIAAAORAKVPAQYTL	600
D	b	177	-----	176
Q	y	601	YSRQRKPELPNNNSNQASMTGAAASGLPPRPLNNSCGTPATTPGVTYPRGRAGVSOS	660
D	b	177	-----GGDAASSWASV-----	187
Q	y	661	OPGSRSTSPSTYLRDQYGGIGNVYRGATGAIPEKASGIPRSTASRSTSPTRSGGLMKR	720
D	b	188	-----	187
Q	y	721	SMYTGAGSRRTPERNNPVPRPAPRLAQSREAEHTTGVDDGDPDYVSGDYMSGMR	780
D	b	188	-----NSEKGSIPR-----	197
Q	y	781	MGRKLMGRDESDIDSEASVCSERSPDSYTRGNKSNYSLSGSHTRLDWQORAPFDI	840
D	b	198	-----SKLSAGSKAHNIAI-----	213
Q	y	841	ETIIFCSTHWSERKGLISLTQYIADGKELTQOOLKCVLDMFRKMFMDTHT----	896
D	b	214	-----SEKKEGKKLLPIVAD-TSLNPIEIKNIGNCMLRLSDASNTWLEIY	260
Q	y	897	SLEFDTVTELLIVANETSRRNGSSCRLRPKXKGTDLNLSHAKIKMTLOYVHEFPPO	956
D	b	261	SLEFVET-----HSSRLS-EWLRLLAKLPARKKAETLPYTKQDIGTLNVILECFNAH	312
Q	y	957	LQKLEFRIISDSTQPTTKTRIALREFTDLANTYCKSDPSPSOACERTVYKLAQL	101
D	b	313	HQLVYVCELMQDPHLMWPVKARVULLEVLTSLDEYTERG-ASTINAEKLTAKIKMLTW	370
Q	y	1017	AADQK-SNELASQARSCLVALYNLTPOMTLLADLPKYVODSARSCJSHMROSQCN	107
D	b	371	ASDPRLSITLPRYHEKAKISMFECVNVADFSALISD-----SEOKMWHOTLOR-----	421
Q	y	1076	SGANPSSSPSSSSSPKPLQSPVGPFPASLSQSHHQJLISSTSS--PR--SROSVEQEL	113
D	b	422	GLENGISSNNIATNS-----GATASRETSNSTFOKESITISFGLPFGARKGTGYNL	472
Q	y	1131	LFSSELDIQHNIOKTSEERHRCFGGQYQOTALPANGFNGLQYHDOGQODSCASLSNSKT	119
D	b	473	-----GSLINNNIALSRLE-----EGSTRLMKEKNL	500
Q	y	1191	QSANQTOSNPPESATKRLDLNERERTONAKSPDIDAKYIVTSINNAENG-----	124
D	b	501	NSIV-TLPPDTELEKIQNVODLLQKRRSENDE--QDEAISISYMMICDGGFGWEQCYA	557
Q	y	1242	ELIILASINMESEVVRVALTLTKDQFVELQTSLTNLGICIKGNCCELPNKJFRSJMRYLL	130
D	b	558	KLLI--NLFE-----ILSKSR-----SENNKXMC-----LRLIG	584
Q	y	1302	NILEAERTDVVIAGLHVLSTKMSRNNKRNHMMHFLLELLIKIIOCYOHSKEALR--DIDS	135
D	b	585	KMCTAQ-----AAKLFQST-----EAAVCCVLDAAVATNDATTALAVED	623
Q	y	1360	MIRPIASLPDLISINIVNPIAAGEPTN-----LCMKILLEVTEHNGSITDAHDIY	1415
D	b	624	CLRTLATHPLPSNINIAKYIL--NOQFIDDERASLVAKMWTIRFEELPRABELNNIVDOI	681
Q	y	1416	FPNIAARSADDTQSMWRKAFFCIYKLVFLGEEKVVKPKLSVL	1457
D	b	682	TPRIKAYOSTSSTVRKTVUYCVJAAVNAVRGEOGMTHFKL	723

```

Db      107 AATCSFTVSKYGIETHSIGEDILVPMASOVAVSTKIMATASLTETFEIYEYQTRQVFT 166
      446 IYDTTLNOSKSDIRSTLCELMVLLFEEMQTKALERNATVLRDLTKKSIDADCDARHS 505
      167 ILS-SFSTSDKSGRRRLALLLEIVISKMSDRIKQIMROI CELIKSAINDAIDASEPBAAG 225
      506 RYAWAFRRHFPPELADQIYGTLDIAQPALEREEREGGGGCTGTGTAPETRTYVRIG 565
      226 RRAFAKIDEMHSEADALYLELDSKQKMLR----- 256
      566 RTGCTLOKTPRMSKISAVDTAAQAKVAQYTLVSRQKRPGLGNNSNGASMTGAASG 625
      257 ----- 256
      626 SLPRRLNNSGCTPATPGSVTPRPRGRAGVSGSQSGSRSTSPSTKLQYGGIGNYR 685
      257 -----GDAASSMASV----- 267
      686 GATGALPKKASGIPRSTASRETSPTSGGGLMKRSMYSTGAGSRRTPERNNPVPSAPA 745
      268 -----NSEKSGIP 276
      746 RLAAQSRBAHTLGVDDGDPDYVSGDYMSGGMRKMGDESDIDSEASVCSER 805
      277 R----- 277
      806 SPDSYTRGNKSNYSLSGSHRLDMSQORAPFDITIIQFCASHTWSEKDGILSTQY 865
      278 ---SKSAGSKAHNNISAI-----SEKKEGKILPPI 306
      866 LADGKELTQOOLKCVLDMFRKFMQDTHT---KYVSLFLDTVELLIVHANEYSRNGSSS 921
      307 VAD-TSLNPEIKNGICNLRLSDASNTWLEIYSIFVRI-----HSSRLS-EWRL 357
      922 CLTRLENKLTDLNMSHAKIWKTLQVHEFFPTQOLKELFRISSTQTPPTKRIAI 981
      358 ALAKLFARKAETLPNTRKQIGHTLNYLLECFNHHDLVYCELMCPHLMVPAKAVVL 417
      982 LRFLTDANTYCKSDPSPDOSQACERTVLKLAOLADOK-SMELRQASSCUVALNLN 1040
      418 LEVYTLSDLEYTEG--ASINAKELKTAIRKMLTWASDPRISILLTHVEKAISMCVN 475
      476 VADSALISDL---DSEQKWHIQTLQIR---NGLNGISSNNIATNS-----G 517
      1101 PPSALQHHQOLISSTS---PR--SROSSVEQELLSSELDIQNIQKTSSEIRHCFGG 1155
      518 ATASRETSNTSFOKESTSFGLPBFGARKGTVNL---GSLINISNNLALSRL- 567
      1156 QYQALAPNGFNHGLQYHDOGQDSCASLSNSKTOSSANTQNTQNPESATMRDLNLER 1215
      568 -----EOSTSRHLEKNLNTV-TLPPDTLEKIONODLLQKA 604
      1216 RTTONAKSPTDKAVITVSIINAENG-----ELIASNLMESEVVRVALLTLQDP 1266
      605 RSSENADE--QESAISSITVMICGGGFGVMEQYAKILL-NLE- 651
      1267 VELLQTSVNLGICIRKGNCELPNKHFRSIRKMLNLLEAHTDVYAGLHVLSKIRSN 1366
      652 -----SENNKMC-----LRILGKCKTAO-----AAKLDPST 678
      1327 KMRHNMHFLLELILKIIQCYHSEKALR--DIDSMIPRIAPSLPLDSINIIVPVATG 1384
      679 -----EMAAKCVLDAAVNTDATTALAVEDCLRTLTHPLPLSINIIVAVIL--N 726
      1385 EFPN---LCAIKILLEVTEHSGSEITDAHLDIVFPNLASADQOSMVKAAVFCIVK 1440
      727 QEPIDDERASIVLKVTRLPELPAELNINVIDITPTIIKAYOSTSIVYKTVVYGLVA 786
      1441 LYPVLGEEKVPKLVNPSKVRLLNVYIEK 1471

```

```

Db      787 MNVRVGEQRMTPHFTKLPRKAMTNLIQVYVNR 817

RESULT 6
S68176
TOG protein - human
C/Species: Homo sapiens (man)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S68176
R/Character: S.; Mazel, M.; Taviaux, S.; Berta, P.; Chow, T.; Larroque, C.
Eur. J. Biochem. 234, 406-413, 1995
A/Title: Characterization of the cDNA and pattern of expression of a new gene over-ex
A/Reference number: S68176; MUID:96128167; PMID:8536682
A/Accession: S68176
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1972 <CHA>
A/Cross-references: UNIPROT:Q14008; UNIPARC:UPI000016787C; EMBL:X92474; NID:91045056;

Query Match      4.3%; Score 328; DB 2; Length 1972;
Best Local Similarity 18.7%; Pred. No. 7, 2e-10;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

      61 SHEKLAQSLAEFSELIKRLGSDPNAVYATVLPVHVIDRIGDSRDIVREKQAQLLRDMER 120
      653 TNEQVQMKLHIVA-LIAQKGNFSTSAQVLDGLVDKIDYK--CGNNAKEMTAIAEA 709
      121 RVLPQALIDKLATSCFKHKNKAVREEFLOTVNALHEVGTQQLSVRYIIPVYCALLGP 180
      710 CMLPMTA--EQVSNMFSQKNPNQSETLNLMSAIKEFGFGLNFKAVISNKTALAA 767
      181 TNNVRAAIQTLVEIKVYGDRLRPDLRM--DDVASKLMLMEQKFDQKQGLLPSA 238
      768 NPAYRTAATLLGLVWMLYVG---PSLRMFEEDEKA-LISQIDAEF--KMGQSGPPAP 820
      239 LKNTNGNV-GLDEADNIGLRERPTMIKRLPSHSAVSSSLRPKNVNDVYGDGAAYTES 297
      821 TRKISGHSTGDEBEGD-----GDEPDGSDNV----- 849
      298 FESSFEVNPQLNIFKAMDNDIYQVLIISDNKADWEKRVDAKKIRALLISYHTPOQ 357
      850 ---DLPRTEL---SDKITSILVSKIGDN--WKIRKEGLDEVAIGI-----NDAK 892
      358 FVAVLKEISSLGVLDLKELELSQVREACTITAMSKTLRNKLDLFCMSIIEHLNLIQ 417
      893 FIQPNIGELPPTA---LKRGLNDS-----NKI-----LVQQTNLTLQ 925
      418 NSAKVIAASTALAKYIYTHAPKLLKIYDTLNLNOSKSDIRSTLCELMVLLFEEMQTK 477
      926 Q---LAVMAGPNIKQHVKNLGIPIITVLGDSKNVRAAALATV----- 965
      478 ALERNATVLRDLTKKSIDADCDARHSHRYAWAFRRHPELADQIYGTLDIAQPALER 537
      966 ---NMAAQGTQKEMLEBEDLSE-----LKKNPFLRQCEILGWL--AEK--- 1005
      538 EREGGGGGGTGTGTAPETRTVSRIGRTPTGLQKPTSMRISAVDTAAOR---AKV 594
      1006 -----LPTLRST-----PTDLILCVPHLYSGLERNGDVVK 1036
      595 RAQYTL-----YSRQRKPLG--PNNSNOASMTGAASGSLPRRLNNSGCTPAT 643
      1037 KADQALPFPFMHGLYKMKMAATGKLKPTSDQ-----VLAMLEKAKVMPPAKPAP-- 1086
      644 PGSVTPRPRGRAGVSGSQSGSRST--SPSTKLQYGGIGNYRGATGALPKKASGIP 699
      1087 PTKATSKPMGGASPAKFPASABEDCISSTPEKPD-----PKA----- 1127
      700 RSTASRETSPTSGGGLMKRSMYSTGAGSRRTPERNNPVPSAPARLLAQSREAEHTLG 759
      1128 -----KADGLSKAKSAQG---KKMPSKTS-----LKEDE----- 1154
      760 VDDGQDPDYV---SGDYMSGGMRKMLM-----GRDE-SDDIDSEASVCS-----ER 805

```

QY 1445 LGEEKYKPLKSLVINSKRVLLNLYIEK 1471
 DB 1027 LGMKMEPHLONLSGKLNLYQVYVNR 1053

RESULT 4

330876
 hypothetical protein R107.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 10-Nov-1995
 C:Accession: S30876
 R:Thomas, K.

Submitted to the EMBL Data Library, July 1992

A:Reference number: S30871

A:Accession: S30876

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1080 <THO>

A:Cross-references: UNIPARC:UPI000017BB14; EMBL:Z14092

C:Genetics:

A:introns: 74/3; 342/1; 560/3; 688/3; 816/3; 963/3

Query Match 7.8%; Score 593.5; DB 2; Length 1080;
 Best Local Similarity 20.7%; Pred. No. 3.4e-25;
 Matches 261; Conservative 208; Mismatches 472; Indels 319; Gaps 39;

QY 288 GDAGAVTMSPESEFEVVPOLNIFPAKMDIDYKQVLVIISDNADMEEKRVALKKIRAL 347
 DB 54 GGTCTVSKDDFLKSEFEVPPKMEISPSDPKEKLDQITETLSKQEDMKNMKLQIISM 113
 QY 348 LILSYHTPOQFVAVOLKELSLSFVDILK---ELRSQVIREACITIAVMSKTLRNKLDAP 404
 DB 114 VV---HGEDVIGREQLLSQVRLTDCLDLSVKDLRSQILREAAITCGFLFKFGTDVROI 170
 QY 405 CMSLIEHLINLQNAKYASASTALAKIITKYTAPLKLYTPTLQNSKIDIRSLIC 464
 DB 171 AERCPSPAPQAVAVATKVMATGAVLTFTVEFIQTKICTGIA-SYSTSKDKNRRLC 229
 QY 465 ELMVLLPEEMOTKALERNATVLRDLTKKSGDADCDARRSHRYAVMAFRHPELADQY 524
 DB 230 ALLEVLVHWNKIRKTYLPQIGELIKALICDADETRVAGKAKSKLDALHSTADLKF 289
 QY 525 GTLDIAAPALEREEREGGGGGGTGTGTAP-ETRTVSRIG-ETPGLTKPTPSMSI 581
 DB 290 ASVDSKQKMLRASDAASSSTINSERGTAPFRSKLSAGSIGIRNAPNISKFLAQRSA 349
 QY 582 SAVPTAAARAKVRAQVTLYSQRKPLGPNNSNQASMTGAAGSLPRRLNSNGCTPA 641
 DB 350 SAIDTKQVTRMAT---SVSRTP---NIRPMTRTILS---KIDTSPGSGKF 390
 QY 642 TTP-GSVTPRP---RGRAGVSGQPGSRSTSPSTKLRDQYGGIGNYRGATGAIIPKA 695
 DB 391 ARPFGALGSTRSSNLRARGSVPTSPGSRNSPPRR---PSATEFPPEM 438
 QY 696 SGIPRSTASSSETSPTRSGGGLMKSMYSTGAGSRRTERRNN---PVPSAPARLLA 749
 DB 439 QRVKSNLNSGNSFVSLSAEATKLOKAMNTAKESLRQPSRNDDEFLPKRPT-PQKATP 497
 QY 750 QSRREHRLGVGDDQGPVYSGDYRSGGMKGRKLMGRDSDSDIDSAASVCSERSFDS 809
 DB 498 QK----- 499
 QY 810 SYTRGNKSNYSLSGSHRLDWSSTORAPFDIETIIQPCASTHSEPKDGLISLTOYLADG 869
 DB 500 -----SALDTSR-----VEEVIKACSTSTANKEKRCIKLAGIYSE- 535
 QY 870 KELTQOOLKCYLDNFRKPMFDHTKYVSLFDYTELILVANETSRNGSSCLTRLENK 929
 DB 536 PNLNAEIKSLGAVLNRLIGESTNQI---VLESISSFVKTHPRLS-DMLKLGKGLPAK 591
 QY 930 LGTDLNLSMHSKIMKTLOVNHVEYPTQOLKELFIISDSOTPTTKRILIALEFVLDLA 989

DB 592 KGAENTLNSKKQISTTISLSSFPDPTQLKSTCELVCDPIHLSPKSRVLLLEYNELL 651
 QY 990 NTYC-KSSDPPSDQACERTVLKLAQLAADQKMEI-RSQARSLVANYLNTPOMTLL 1047
 DB 652 GKIMERGSSFTYKMKMA---TILKMFWMADQRNEQLITPGEKVLCSLFALNNADEFAL 708
 QY 1048 LADLPKYQDSARSCISHMRRO-----SQCNSGANSPPSSPLS-----SSSPK 1092
 DB 709 FNDFNPDYRDVAAYKVLQSHGHQHVPOQDAVEBACVATISTATQAEDEFEVRNLDMT 768
 QY 1093 PLOSPVGPFPALSHHQLQSSSTPSRSPSSVEQL--FSSELDIOHNIQITSEIR 1150
 DB 769 FVKSPT-----RAISSGPKR-----VDAEPLRLPSSKMSQHRDEELS----- 807
 QY 1151 HCFGQYOTALAPNGFNGLQYHDQGDSCASLSSNKTQSANTQSPESATWRLD 1210
 DB 808 -----FN-----ESTDRKLNSTHLLIDTISB----- 829
 QY 1211 NLERERTONAKSPTDKAVITVSINN-AENGELILASNMESVVRVALTLYDQVEL 1269
 DB 830 -----QSKYVASKLAQISGDWGAQOYEGILS----- 855
 QY 1270 LQSTLTNLGICIKGNGELPKHFRSITMRMLNLEHMDV-VIAGLHYTKIMRSKYM 1328
 DB 856 IQTWLCE-----GSFTLWQONPAKLLIAVFDVLSKESBANKKVALRVLTCKTQSOAS 908
 QY 1329 RHNMMHFLLELLIKIIOYHKEALRDI--DSMIPRIAPSLPLDSINI----- 1376
 DB 909 R-LFDSTEWALICVLDAAANSQDGTNNVATDCLKTALHTPLAKYVINSQLLNEKA 966
 QY 1377 -----VNPVIATGEFPTNI-CAIKILLEVEHHSSEITDAH 1411
 DB 967 QEPKASIVLKMRTFLFEGQLADELSPVV-----DILAFCVIAKTYIQTSQFINEIF-- 1017
 QY 1412 LDIVFPLASADPTQSMRKAAPFCIVKYFYVIGEEKVVPKSLVNSPKRYLLNLYIEK 1471
 DB 1018 -----TLQSYSPSSAVKRTAYVCLVAVNKKLMKMTMBHQLNLSGKLNLYQVYVNR 1070

RESULT 5

C88550
 protein ZC84.3 [imported] - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
 C:Accession: C88550

R:anonymous, The C. elegans Sequencing Consortium.
 Science 283, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 A:Reference number: A75000; MUID:9069613; PMID:9851916
 A:Note: see websites genome.watli.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.e

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a
 A:Accession: C88550

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-826 <STO>

A:Cross-references: UNIPROT:Q03609; UNIPARC:UPI000013BAAC; GB:chr_III; PIDN:CAA79568.1

C:Genetics:

A:Gene: ZC84.3

A:Map position: 3

Query Match 5.1%; Score 389; DB 2; Length 826;
 Best Local Similarity 17.1%; Pred. No. 6.8e-14;
 Matches 211; Conservative 203; Mismatches 371; Indels 446; Gaps 35;

QY 273 VSSSLRPKPVNVVDTGAGVNTMESPESEFEVVPOLNIFPAKMDIDYKQVLVIISKNA 332
 DB 1 WTSRQPKSGVY-----SVSKADFTKIKFEDVPPKPTISAVDRKRFDAVRILISNSSE 53
 QY 333 DMEKRYDALKKIRALLILSYH-----TOPOFVA--VOLKELSLSFVDILKEEBSQVIRE 385
 DB 54 DMNRQGTQKTVASIVY---HGEKVVDPRPTMIAHLQV-----LGCELANVKDLRSQVIRE 106
 QY 386 ACTTIYMSKTLNKKLDAPFWSLIEHLINLQNSAAYIASASTIALKYIITKTHAPKLLK 445

```

0Y 930 LGTDLNMSKXMTLOVNEHYEPTQOLKELPRISDSQTPTTIRIILNFDLDA 969
Db 775 MGADLLGSVQAKQKALDVTRSPENDLQNFILMRFTYDQOTQTSLKVAIIXITFLA 934
0Y 990 NTYCKSDPSPD--QSOACERTVLKLAQLAADQKSMELRSQARSCLVALYNLTPOMTLL 1047
Db 835 ----KQMD-PDDFINSETRLAVSRVITWTTEPKSSDVRKAQSVLLISFELNPEFTML 889
0Y 1048 LADLPKYQDABARCSITHMRQSQSCNSGANSFSSSPFLSSSPK-----PLQSSVG 1100
Db 890 LGALPKPTQDQATKYLHNHMLR---NTGNGQSSMGSPLTPTPRSPAMNSPLTSPPT 945
0Y 1101 PFASIQSHHQLSISSTSPRSRSVE-----QELPFSSBLDIOHNIQKTSSE---IRH 1151
Db 946 SQNTLSPAPFYDTENNMSSEDIYSLRGVTAIQFSPRSQEDNNEPKRDSKDDDDSM 1005
0Y 1152 CFGGQYQYALAPNGENGHLYQHDOQGOQDSCASLSSNKTQSSANTTQSN----TPESATM 1207
Db 1006 CGG-----PGMSDDPRAGDQATDSQOTALDNKSLHSMPTHSSPRSDROYNPYNSDSISPF 1061
0Y 1208 RLDNLERERTQNAKSPPTDQAKVITVSINNAENGELILASUNLMSVYVRVALITKQPV 1267
Db 1062 NKSALKKEMFPDDDAQPPDD-----LSLDHSD---LVAEILKELS-----NHNERY 1104
0Y 1268 ELLOTSLTNLIDICIKGNCCELPNKHFRSIMSLNITIEAHTDVIVAGLHYLSKIMSNNK 1327
Db 1105 EERKIALYELMKLTQOESFSYWDHFKITLLLETJGODEPTTRIALALKYLREILNHOP 1164
0Y 1328 MRHNHMFLELILKLIIOCYQH-SKEALRDIIDSMIPRIAPSLPLDLSINTVNPVATGEF 1386
Db 1165 AR--FKQYAEITVWKTELAHNDPKHEVVRSAEBAASVLAISISPEQCCKVLCPIQYADY 1222
0Y 1387 PTNLCAKYLLEVTBHHGSEITDAHLDIVFENLARADDTQSMYRKAARFIVLYVLG 1446
Db 1223 PINLAIMQKQVIERVSKETLNLLEPIMGLLIQGYDNSSSVYKCKAVCFCLVAHVAVIG 1282
0Y 1447 EEKXPKTSLVNPSPKVRLLNYIEKORNCISGSSSTKNSSAAS 1490
Db 1283 DE-LKPHLSQLTGSKMKLNLNLYI-RAQOTSGGADPTTDVSGQS 1324

```

```

RESULT 3
E88546
Protein R107.6 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: E88546
R/anonymous: The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: E88546
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-1063 <STO>
A/Cross-references: UNIPROT:P32744; UNIPARC:UPI0000610FC; GB:chr_III; PIDN:CAAT8472.1;
C/Genetics:
A/Gene: R107.6
A/Map position: 3

```

```

Query Match 8.0%; Score 611; DB 2; Length 1063;
Local Similarity 20.9%; Pred. No. 3,4e-26;
Match 257; Conservative 213; Mismatches 487; Indels 270; Gaps 36

Qy 288 GDAGAVTSESFSESEVVPOLNIFAKQDDIDYKQVLVIISDKADWEKRYDALKKIPAL 347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 54 GGTCTVSKDDPLKSEFVVPKMTSSBSDPKELDQTIETLSKGGEDANKKRNKKLQIISM 113
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 348 LILSYHTPOFVAVALCKELSLSPVDILK---EELRSQVIREACTITAYMSKTLRNKLDAF 404
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db	11	VV---H8EDV1GRRQQLLSQVLRLTDCLDLSVKDLRSQ1LEEAATCGFLFKRFGTDVRO1	170
OY	405	CWS1LEHLINLIQNSAVIVASASTIAIKTIIKTHAPKLIKITYDTLNOQSKNDIBSTLC	464
Db	171	AERCLPAPAFQAVAVSTYVMATCGAVLTFIVEF1QTKOIFCTIA--SYSTRKDNQORQLC	229
OY	465	ELMWLLPEEQOTALEBNATVLRDLTKKSI:GDADCDARHRSRYAVYMAFRBHFPELAOIY	524
Db	230	ALLETVEHNMKEIKRTVLPOIGELIQA:CDADPEFRVAGRAKAFSLDIALHSEADKLF	289
OY	525	GTULIDIAORALEREEREGGGGGGTGTGTGNAP--ETRRIVSIG--RTPTLOKPPRMSRI	581
Db	290	ASVSSSQKMLRADAASSSTSI:NSEKGTAPRPSKLSAGSIGIRNAPN1SKEFLAQRSA	349
OY	582	SAVDTAARQAPKAPVQYTLVSROKPELGPNNNSQASMTGAAGSLRPLPLNSGCTPA	641
Db	350	SAIDTKQVTRMAT-----SVSRTP-----NIRPWTTRTLS-----KIDTSPGSKPF	390
OY	642	TTP--G5VTRP-----RGRAGVQSQOPGSRSTSPSTLRQOYGIGNYGATGAIPIKA	695
Db	391	ARPVGALGRTSSNLARASVPTSPQSRKNGSPRR-----PSATEAPFAEM	438
OY	696	SGIPRSPASSRETPSTRSGGLMKRSMYSYAGASRRTPERRNN-----PYRSPAPARILA	749
Db	439	QKVSNNGSNSFYVSSLSAEATKLOKXMTNAXKSLRSPSNDDEFLPRPFT--POKATP	497
OY	750	QSRBAEHTLGVGDGDPODYVSGDYMRSGNRMGRKLMGRDESDDIDSEASSVCSERSFDS	809
Db	498	QK-----	499
OY	810	SYTRGKSNYSLSGSHRDLMWSTORAFDIEITIOFCASTHMERKDGILSLTQVLADG	869
Db	500	-----SALDTSR-----VEEYTRCSSTSAEKEGKIMLAGIYSE--	535
OY	870	KELTQOOLKCVLDMFRQMFMDTKVYSLFEDYTELILVHANETSIRNGSSCCTRLPFXK	929
Db	536	PNLSNAIKS1GAVLNRILGESTNOI---VLES1SSGVKTHHRLS--DWLKG1GKFLAK	591
OY	930	LGTDLNLSMHSKIWKTLQVNVHEYFPTOLOKELFR1ISDSTQRTTKYRIALRFLTDLA	989
Db	592	KGAMTILNSKKQI:STTICILSSFPDPTLOKSTCEIJCDBD1HLMSPKSRVLLLEYLWELL	651
OY	990	NTYC--KSDPSPDOSQACERTVLTQAOLADQKSMEL--RQARSCVALVNLNTPOMTLL	1047
Db	652	GKUYMERSSFPYTKEMKA---TILKMF5MADQREOELITPHEKVLCSLFPALNNADPSAL	708
OY	1048	LADLPKYODSARSC1SHSHRRQO-----SOSCSGANSPESSPLS-----SSSPK	1092
Db	709	FNDNPNYRDMAYVVLVUSHGDQHVPOODAVSEACVRAIT1STTAAGIDEDFVSRNIDMT	768
OY	1093	PLQSPSVGPFPASLOSHHNOLISISTSPRSKOSVYEOEL--FSSELDIOHNIQKTSSEIR	1150
Db	769	PVKSPPSR-----RAISSGFRK-----VDAPRLPLPSEKMSQHRDEBLS---	807
OY	1151	HCFFGQYQALAPNGFNHLLQYHDQGOODSCASISSNSKTSQSSANTQSTPESATYRLD	1210
Db	808	-----FN-----ESFDRKJLKNSTTHLIDDTSE-----	829
OY	1211	NLERERTONAKSPSTDAKIVTYS1NM--AENGELILANSMESEVVRVALLTIDQOVEL	1269
Db	830	-----QSKTYVASKIAQJISGDGAQOQYBLS-----	855
OY	1270	LOTSLTNAGICIKNGCNCELPNKFRPS1MRMLN1L1EAHTDV--V1AGLHYLSK1TRMSNM	1328
Db	856	IQTMLCE-----GSFTLMEQNQAPKLLIAVFDVLSSSESDANKVVALRVLTWKCTSOAS	908
OY	1329	RHNTMHLELILKTIQOQSKKALDI--DSK1P1AASLP1B1DLS1NVN1P1Y--ATG	1384
Db	909	R--LFDSTEWAIKCVLDAAVNSQGTGNVAVADCLKTATLPHLPLAKVNV1SOL1NEEKA	966
OY	1385	EPFNNLCAIKLLEVENHHSSETDALD1VFPRLASADTOSMKAARVFC1VKLXYFV	1444
Db	967	QEPASIVLTKMTRTLFEGLOADELSFVVDLDLAPCV1AS1YSPSSAVKRTAYCVLAAVNVK	1026

LENGTH: 154
 TYPE: DNA
 ORGANISM: Drosophila melanogaster
 US-09-270-767-30401

Query Match
 Best Local Similarity 99.3%; Pred. No. 4.9e-34;
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2154 GGAATATCCCGAAGCAGACCCAGCTCTAGAGAAAGAGTCAACTAGAGTGGTGGC 2213
 1 GGAATATCCCGAAGCAGACCCAGCTCTAGAGAAAGAGTCAACTAGAGTGGTGGC 60
 2214 TTGATGAAGACGAGTATGTAATCTCTACAGTGGGGGCTCCAGCTACCGCCGAGAAAC 2273
 61 TTGATGAAGACGAGTATGTAATCTCTACAGTGGGGGCTCCAGCTACCGCCGAGAAAC 120
 2274 AACCCAGTAAAGCAGTACCGCGCGGCGAGCAGC 2304
 121 AACCCAGTAAAGCAGTACCGCGCGGCGAGCAGC 151

RESULT 8
 US-10-104-047-517
 Sequence 517, Application US/10104047
 Patent No. 6943241
 GENERAL INFORMATION:
 APPLICANT: HELIX RESEARCH INSTITUTE
 TITLE OF INVENTION: NO. 6943241el full length cdna
 FILE REFERENCE: H1-A0105
 CURRENT APPLICATION NUMBER: US/10/104.047
 PRIOR FILING DATE: 2002-03-25
 PRIOR APPLICATION NUMBER:
 PRIOR FILING DATE:
 NUMBER OF SEQ ID NOS: 4096
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 517
 LENGTH: 2524
 TYPE: DNA
 ORGANISM: Homo sapiens
 US-10-104-047-517

Query Match
 Best Local Similarity 46.1%; Pred. No. 6.2e-31;
 Matches 769; Conservative 0; Mismatches 852; Indels 45; Gaps 7;

77 GAAAGCCAGGACCTGATGCTTTATTCAGCAAAATGCCAAGGCGAAGCATGCGTGA 136
 357 GAGAGCCCGCAGCATGAGTACTTCTGCGCCAGGTGAGAGAGCGTGGGCGC 416
 137 GGTACAGCTGCGGAGATCTGTGACATTCCTTAGCG--ACGACCAAACTCAATTGT 193
 417 GCTGAGGTGCGGCGAGAGCTCTGCTACCTTGGGCCCCCGGCCCATCTCGACCT 476
 194 CTGACGAGATAGGATCTCT-----TATTGAGGTTTGTGTCATGCTGACGG 244
 477 GAGAGGAGACCTGGCGCGCTAGGCAAGACAGTGCAGCGCTCACCGGTGGGTTCT 536
 245 CAGCAGCTTAAGATTGCAAAAAGTCCCTGAGGCGCTTCTGGAGTATTAAGCAT 304
 537 GAGCAACTACCGGATCATTAATGGATGGAATTTTAAGTGCCTTTGTGACAGATT 596
 305 GGGCAGCATTTTAATGATATACAGGCTACCGTCTGCGACATGATGATGAGTGG 364
 597 ATCAACACGCTTAAATCTATGATGCAATGTTATGTATGCTTAAATAGACAGATT 656
 365 AGACAGAGGAGACAGTCCGCGAGAGGCGCAACTTCTGCTCGGAGCCTCATGAGCA 424
 657 AGATGCCAAAGACAGATGATGAGTGAAGCTCAGACTCTGATATTGAAGTTAATGATCA 716
 425 CAGAGTGTTCGCGCCAGGCGGTGATGCAAGCTGCGCAGTACGCTTCAAGACAA 484
 717 AGTAG-----CAGCAGTATGATGATTTGGAGAGCGTTGCTTGTGTTAAACAAA 770

485 GAAAGCCAGGTCGCGAGAGTCTCTTCAAGCATGATGAAAGCTCTTCATGATGACG 544
 771 GAAATTCGATCTGAGAAAGGCGTGTCTGTGTCTTATGTAACCTTAAACATTTTGG 830
 545 CAGCCAGAGCTTACTGTTCGCTCTATATACACCACTTTGTGACCTTTCGAGATCC 604
 831 GGTCTAGCCATAGTCAACAGCAATGATACCATTTGTGTATCTGTTTGGAGACTC 890
 605 CACAGTTAATGTAGGAGGCGGCATCAAGCGTATGTAATCTCAAGCATGTAG 664
 891 CACAGTACGATGAGAGTGTGCAATTTGGCTATAGTGAAGATTATTAAGCATGTGG 950
 665 GGAATGATTTGGCCAGACCTCCGTCATGAGAGATTTCTCTCC-----TCGAAATTGG 720
 951 AGAAAGATGAGATGATCTTTATTAAGAGAGAAATTCCTCTCTATATTAAGAAATGAT 1010
 721 CTATGTGAGCAAAAGTTGACCAAGTCAAAACAGAGGCTCTACTGCTACCTTCAGCC 780
 1011 ATTTGCCAAATTTGATGAGTCAAAAGTTCAAGCGGTATGATTTGAGTGTCTGCAAGA 1070
 781 TTAATAACAGAAATGGAATGAGTGGCTTTGACAGGCGGACAAATTTGGTTGAGG 840
 1071 TAAAGCTTGATGATGAGATGAGATGAGATGAGAAATGAGCATCATCAGTGCATGAC 1130
 841 AGCAGCCACAGGATGATTAAGCGGCCACTACACTCGGCCGTTTCG-----887
 1131 CTTCAGATTCCTGACACTTAAACATCCGAAATCTCCCAAGTGCAGAGAACTTGG 1190
 888 -TCATCACTCGCCCAAAACCAGATGAAAGATGAGTACCGGTGATGCGCGCGTAAAC 946
 1191 TTCAGCAGTGGCCCTTAAGGTGAGAGTGTCTTCAAGAAAGAGAGTGTGAGCAGTTGA 1250
 947 CAGGATCTTTGAAATCTAGTGTGAGTGTGTCCTCCCAATGAACTCTTCCACGCTAA 1006
 1251 TGAAGATATTTATTAAGACCTTTTACAGATGTCCTTCAATTCAGATTTATTTAGTGC 1310
 1007 GGAATGAGCATATCTACAGCAAGTACTAGTATCATGATGATTAATAAGCAGACTG 1066
 1311 AGAATCGAAGAAACATTAATAATTAATCAAGGAAATTTGTCAATATTAATCACTG 1370
 1067 GGAAGAACTGTGATGCTCTCAAGAGATCAGGCGATTCATTTCTCAGCTATCAAC 1126
 1371 GATACAGCGGCAATGACATGAAGAAATTCATCATCTCTGTT-----GCTGAGC 1424
 1127 TCAGCGCAGTGTGTGCTGTACGCTAAAGAAATTTGCTTAACTTCTGTCGATCT 1186
 1425 TGCACAGTATGATGCTTTTTCACATTTACATTTGATGAGACCTTAAACTTTTC 1484
 1187 CAGAGAGAACTACATCAGATGATCGCGGAGGCGTGCATCACCATGCGCTTACATGTC 1246
 1485 AGCTAAGATCTTAATATCCAGGTGTGAGAGAGCTTGTATTAATCTGTAACCTTTC 1544
 1247 TAAAGCTGAGAAATTAATAGATCCTTCTGAGAGATTTGAGACCTGATTA 1306
 1545 AACAGTTTGGGAAACAAAGTTGATCATGCGCTGAAGCATGTACATTTTAA 1604
 1307 TTTAATACGAACAGGCGCAAGTCTTTCATCCGTTCCACATATGCTTGAAGTAT 1366
 1605 TCTCGTCCCAATAGTCAAAAGTCAATGCAACTTCTGATGTGACGAATCAGATTAT 1664
 1367 CATTAAGTATACATGACAGCAAGTGTCTCAAGATCTACAGACACTGTAATCAT 1426
 1665 CATTCGGCATATCATATGTAACCACTTATACCTTTAATACAGCAATTTGCA---CATC 1721
 1427 AAAGTGAAGACATTAAGTTCACATGTGTGAGCTGATGTGTCTCTTCGAGAGTG 1486
 1722 AAATCAGTTCCTCGAGAGAGAGTTCATTTAATTTTAAGTTATATGTTGAAGAGTG 1781
 1487 GGAAGCAAGGCTTGAAGAAAGCAACCGTACTAAGGACACTTAAATAATCAT 1546
 1782 GCAAGCTCATTTATGAAAGACATGACAGCGCTTGTGTTGAATATTAATAAGGAAT 1841

Query Match 2.3%; Score 116.6; DB 3; Length 1908;
Best Local Similarity 48.7%; Pred. No. 1.7e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

US-10-104-047-113
Sequence: 113, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241e1 full length cDNA
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104.047
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 113
LENGTH: 1908
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-113

1547 TGGGATGCAAGCTGCGATGCAAGCCCGCATTCAGATACCCCTTATGGCTTTCAGCG 1606
1842 TCATGATGCTGACCCCTGAGGCCAGAGTGAGCAAGAAAGACATACATGCGCTTTAGAAA 1901
1607 TCACCTTCCAGAGCTGGCGGATGCAATATATGAAACATTAGACATAGCTGCCAGCGGC 1666
1902 CCACCTTCTGCTGTAAGCTGAAACATTATATATTCCTTGAAGCCATCTTATCCAGAAAG 1961
1667 ATTGAAAGGAGACGAGAGCGCGCGAGAGAGAACTGGTAC 1711
1962 TCTTCAAACTTACTTAACAGTTCTGCGAGTGTGACATCTTCC 2006

Query Match 2.3%; Score 116.6; DB 3; Length 1908;
Best Local Similarity 48.7%; Pred. No. 1.7e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

927 GGTGATGCGCGCGCGGATGCAAGTGAATCTTTCGATCTTTCAGAGTGGTCCGCA 986
353 GGAAGTGTGAGAGAGTGTGATGAGATGATTTTAAAGCTTTTACAGATGTCCTTCT 412
987 TTGAACATCTTCCAGCTTACAGATGAGATATCTTCAAGCAAGTACTGTATCATC 1046
413 ATTGAGATTTATTTAGTGCAGAACTCGAAGAAACATTAAATAGGAAATTTTG 472
1047 AGTATATAAAGCAGACTGGGAGAAAGTGTGCTCTCAAGATCAGGCAATG 1106
473 TCAGATGATTAACATGACTGGGATCAGCTGCCAATGACATGAAATTCATCACTG 532
1107 CTCATTTCCAGCTTACACTCAGCGGAGTGTGCTGTACAGCTTAAAGAAATGTCG 1166
533 CTGTTGCTGAGAG-----CTGCAAGATATGATGCTTTTCAACATTTAGATGTTG 586
1167 TTAAGCTTGTGAGATCTCTAAGAGAGAACTACATACAGGTGATCCGCGAGCGTGC 1226
587 GATGAGACCTTAACCTTCAAGCTAGATCTTATAGTCCAGGTGTTAGAGAACTTGT 646
1227 ATCAACATGCGCTCATGCTTAAGAGCGTGAAGAAATAAGATGCTTGTGCTGAGC 1286
647 ATTACTGATGCCACCTTTCACAGCTTTGGGAAACAAAGTTGATCATGCGCGTGAAGC 706
1287 ATTTTGGAGCACTGATTAATTTTATACAGAACAGCGGAAAGGTGATTCATCCGCTTCC 1346
707 ATTGATCTACACCTTTTAACTGCTGCCCAATAGTGCAGAAAGTGCATGCAACTTCTGA 766
1347 ACAATAGCTTGAAGTATATCTTAAATATACATGACCAACAGCTGCTCAAGATCTAC 1406
767 TGTGAGGATCAAGATTAATATCATGCGATCTCATGTAACCAAGCTTATACCTTTAATA 826
1407 ACAAGACCTGGAATCAATCAAGTGAAGACATAGAGTCCACACTGTGAGGCTGATG 1466
827 ACAGCAATTCGA-----CATCAAAATCAGTTCCCGTAGAGAGACGTTCAATTGAATTTTA 883

Query Match 2.3%; Score 116.6; DB 3; Length 6445;
Best Local Similarity 48.7%; Pred. No. 3.9e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

US-09-620-312D-324
Sequence: 324, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620.312D
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_Fl_genes Version 1.0
SEQ ID NO 324
LENGTH: 6445
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4051)
US-09-620-312D-324

Query Match 2.3%; Score 116.6; DB 3; Length 6445;
Best Local Similarity 48.7%; Pred. No. 3.9e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

927 GGTGATGCGCGCGCGGATGCAAGTGAATCTTTCGATCTTTCAGAGTGGTCCGCA 986
362 GGAAGTGTGAGAGAGTGTGATGAGATTTTAAAGCTTTTACAGATGTCCTTCT 421
987 TTGAACATCTTCCAGCTTACAGATGAGATCTTACAGCAAGTACTAGTATCATC 1046

```

Db      422 ATTGAGATTATTATGTCGAGAGACTCGAAGAAACATTAAATCAAGGAAATTTTG 481
QY      1047 AGTGTAAAAACGAGACTGGGAGAAACGTGTGATGCTCTGAGAGATCAGGCGATG 1106
Db      482 TCAGATGATTAACATGACTGGGATCAGCGTGCATATGCAAGAAATTCATGATCAGT 541
QY      1107 CTCATTCGAGTATCAGACTGAGCGGAGTTGTCGTGTCAGCTTAAGAAATTTGCG 1166
Db      542 CTGTGCTGTGAG-----CTGCACAGATGATGCTTTTTCACATTTAATCGATTTGTTG 595
QY      1167 TTAAGCTTCGTGAGATCCTCAAGAGAGAACTACAGATCAGAGTATCCGAGGCGTGC 1226
Db      596 GATGAGACCTTAATCTTACGCTAAGATCTTAAGATCCAGGTGTTAGAGAACTTGT 655
QY      1227 ATCAGCATGCGCTACATGCTTAAGACGCTGAGAAATTAACATAGATGCTTGTGAGC 1286
Db      656 ATTACTGAGCCCACTTTCAACAGTTTGGGAAACAGTTTGAATCATGCGGTGAGCC 715
QY      1287 ATTTGGAGCAGCTGATTAATTTAATACAGAAACCGGAGAGTTCATTCGCTTCC 1346
Db      716 ATTGATCTTACACTTTTAAATCTGTCCTCCCAATGATGCAAAAGTCATGGCACTTGA 775
QY      1347 ACAATAGCTGAGATATATCATTTAATAGATATACATGACCAAGCTGCTCAAGATCTAC 1406
Db      776 TGTGACAGATCAATTTATCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 835
QY      1407 ACAGACACTCTGAATCAATCAAGTCGAAGAGACATAGTCCAGACTGTGAGCTGATG 1466
Db      836 ACAAGCAATGCA---CATCAAAATCAAGTCCCGTGAAGAGACGTTTCAATTTGAATTTTA 892
QY      1467 GTGTGCTCTCTGAGAGTGGCAGACGAGGCGTGGAGAAAGATGCCACCTGACTAAG 1526
Db      893 GATTATTTGTTGCAAGATGGCAGACTCATTTGGAAGACATGACGCGTCTTGTGTT 952
QY      1527 GACACCTTAAAAAATCCATTTGGGATGAGAGACTGCGATGCAAGCCGCCATTCAGATAC 1586
Db      953 GAAATCTTAAAAAGGAAATTCATGATGCTGACCTGAGGCGCAAGTGGAGGCAAGAAAG 1012
QY      1587 GCCATTTGGGCTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAAAATATATGAACATTA 1646
Db      1013 ACATACATGGGTCTTAAAGAACCACTTCTGCTGAGAGCTGAATATATATATATATATAT 1072
QY      1647 GACATAGCTGCCCGCGGCTTGAAGAGGAGACGAGGCGGCGGAGAGAGAGACT 1706
Db      1073 GAGCCACTTATACAGAGAGTCTTCAACTTAAGAAGTCTGCGAGTGTAGCATCT 1132
QY      1707 GTTAC 1711
Db      1133 CTTCC 1137

```

RESULT 11

US-09-620-312D-323

Sequence 323, Application US/09620312D

Patent No. 6569662

GENERAL INFORMATION:

```

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jien-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghaast
APPLICANT: Dermanac, Radoje T.

```

```

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725.
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PL_FL_genes Version 1.0
; SEQ ID NO 323
; LENGTH: 6487
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(4093)
US-09-620-312D-323

```

```

Query Match      2.3%; Score 116.6; DB 3; Length 6487;
Best Local Similarity 48.7%; Pred. No. 4e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

QY      927 GTGTATGCGGCGCGCTAACCATGAAATCTTGAATCTAGCTTTGAGGTGTCGCCAA 986
Db      362 GAGAGTCTGAGACAGATGATGAAGATTTATATAAGCTTTTACAGATGTCCTTCT 421
QY      987 TTGAATCTTCCAGCTAAGACATGAGAGATATCTACAGCAAGTACTGATCATC 1046
Db      422 ATTGAGATTTATTTAGTGTGAGAACTGAGAAACATTAATTAATCAAGGAAATTTTG 481
QY      1047 AGGATTAACCAACGACATGAGAGAAACGATGATGCTCTCAAGAGATCAGGCGATG 1106
Db      482 TCAGATGATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 541
QY      1107 CTCATTTCACTATACACTGACGCGAGTTTGTGCTGTACAGCTTAAGAAATTTGCG 1166
Db      542 CTGTGCTGTGAG-----CTGCACAGATGATGCTTTTTCACATTTAATGATTTGTTG 595
QY      1167 TTAAGCTTGTGAGACTCTCAAGAGAGACTACATCAAGTATCCGAGGCGTGC 1226
Db      596 GATGAGCACTTAATCTTTCAGCTAAGATCTTAAGATCCAGGTGTTAGAGAGCTTGT 655
QY      1227 ATCAGCATGCGCTACATGCTTAAGACGCTGAGAAATTAACATAGATGCTTGTGAGC 1286
Db      656 ATTACTGAGCCCACTTTCAACAGTTTGGAGAAACAGTTTATATGAGCGCTGAAGCC 715
QY      1287 ATTTGGAGCAGCTGATTAATTTAATACAGAAACGCGGAGAGTCAATGATCCGCTTCC 1346
Db      716 ATTGATCTTACACTTTTAAATCTGTCCTCCCAATGATGCAAAAGTCAATGCAATTTCTGGA 775
QY      1347 ACAATAGCTGAGATATCATTTAATAGATATACATGACCAAGCTGCTCAAGATCTAC 1406
Db      776 TGTGACAGATCAATTTAATCAATTCGCAATCTCATGATCCAGACTTAATACCTTTAATA 835
QY      1407 ACAGACACTCTGATCAATCAAGTGAAGAGACATPAAGTCCACATGATGTGAGCTGATG 1466
Db      836 ACAAGCAATGCA---CATCAAAATCAAGTCCCGTGAAGAGAGCTTCAATTTGAATTTTA 892
QY      1467 GTGTGCTCTCTGAGAGTGGCAGACGAGGCGTGGAGAAAGATGCCACCTGACTAAG 1526
Db      893 GATTATTTGTTCAAGAGTGGCAGACTCTTCAATTTGAAGAGACATGACGCGCTTGTGTT 952
QY      1527 GACACCTTAAAAAATCCATTTGGCGATGAGACTGAGATGACAGCGGCCATTTCAATAC 1586
Db      953 GAAATCTTAAAAAGGAAATTCATGATGCTGAGCGCTGAGCGCAAGTGGAGGCAAGAAAG 1012
QY      1587 GCCATTTGGGCTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAATATATATAGAAACATTA 1646
Db      1013 ACATACATGGGTCTTGAAGAACCACTTCTGCTGAGGCTGAGAAACATTTATATATTCCTTT 1072
QY      1647 GACATAGCTGCCCGCGGCTTGAAGAGGAGACGAGGCGGCGGAGAGAGAGACT 1706

```

Db 1073 GAGCATTATATGAGAGAGCTTCAACTTACTTAAGAGCTTCTGGAGGTAGCATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTC 1137

RESULT 12

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F18
US-08-232-463-14

Query Match 0.9%, Score 47.2; DB 2; Length 7218;
Best Local Similarity 6.5%; Pred. No. 0.023; Indels 0; Gaps 0;
Matches 16; Conservative 141; Mismatches 89;

QY 4865 CTCGCAATCTCACTTCTCATCATGCTTGTAGCTTTTACGAGCTCCTAATAA 4924
Db 1218 YY 1277
QY 4925 CGCGATTCGTATTAATTGTAAGCTATATACCGCTCTTAATGAACTCAGACGTGTG 4984
Db 1278 YY 1337
QY 4985 CTTTCACTTGTTCAGGCGACTGTATCCTTAATGTTGTTCCGCTCACTAATAT 5044
Db 1338 YY 1397

QY 5045 CTAAATCTATTATTAATAGATTCCTCTTACAACTGACATTTTGTAAATAT 5104
Db 1398 YY 1457
QY 5105 AAATAC 5110
Db 1458 AACTAC 1463

RESULT 13

US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHIEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F18
US-08-232-463-14

Query Match 0.9%, Score 43.8; DB 2; Length 7218;
Best Local Similarity 8.9%; Pred. No. 0.24; Indels 177; Gaps 0;
Matches 36; Conservative 190; Mismatches 177;

QY 1337 ATCCGCTTCACATAGCTCTGAAGTATATCATTAAGTATACATGACCAAGCTGT 1396
Db 1517 ATTCAGTTTCAAAAAGGAGATGATGACATCTGAATTAATCTATGCAAGTAGTT 1458
QY 1397 CAGATCTACACAGACATCTGAAATCAATCAAGTGAAGAGACATAGCTCAGCTGTG 1456
Db 1457 AAAGAGATGAAGAAATTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1398
QY 1457 TGAGATGATGAGCTCTCTGAGAGAGGACGAGAGGCTGTGAAGAGATCCAC 1516
Db 1397 RRR 1338

```

Qy 1517 CGTACTAAGGACACCTTAAAAATTCATTGGCGATGACATGACATGACGCCGCCA 1576
Db 1337 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1278
Qy 1577 TTCGATAGCGCTTATGGCTTTTCAGGCGTCACTTTCAGAGCTGGCGGATCAATATA 1636
Db 1277 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1218
Qy 1637 TGGACATTAAGCATAGCTGCCAGCGGCATTAGAAAGGACGAGCGCGCGAG 1696
Db 1217 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1158
Qy 1697 AGGAGAACTGTCTACTGGGACTGGGACTGCACCTGAAAGCA 1739
Db 1157 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1115

```

RESULT 14

```

5244792-1/c
; Patent No. 5244792
; APPLICANT: BURKE, RAE L.; PACHL, CAROL; VALENZUELA, PABLO D.T.
; TITLE OF INVENTION: EXPRESSION OF RECOMBINANT GLYOPROTEIN
; FROM HERPES SIMPLEX VIRUS
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/587,179
; FILING DATE: 20-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 921,730
; FILING DATE: 20-OCT-1986
; APPLICATION NUMBER: 597,784
; FILING DATE: 06-APR-1984
; SEQ ID NO:1:
; LENGTH: 3472
5244792-1

```

```

Query Match 0.8%; Score 42.8; DB 9; Length 3472;
Best Local Similarity 54.4%; Pred. No. 0.29;
Matches 86; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

```

```

Qy 3068 CCAGAGCCAGGCTCGCCAGCGGACGCTCTAAAGCTGCCCGCGGATCGAA 3127
Db 1508 CGAGAGGAGTACTGGGTGAGGTGAGTGAAGTGTGAGATGCGTCCGAGAGAA 1449
Qy 3128 GTGATGAGCTGGCTCCAGGCGGAGCTGCTAGTGCCCTGTATACTGAATAC 3187
Db 1448 GCGAAGGAGCCGCGGCTACTCGGCGGAGCATCTGTCACCTCTGCACTTGTCAT 1389
Qy 3188 CCGCAATGACCTTTTACTGGCGGACCTGCCAAGG 3225
Db 1388 GTGCAAGACCGCGGTGCTTCCGACCCAGTCCAGG 1351

```

RESULT 15

```

US-09-561-077C-25
; Sequence 25, Application US/09561077C
; Patent No. 6706501
; GENERAL INFORMATION:
; APPLICANT: Rosson, Reinhardt D.
; APPLICANT: Deng, Ming-de
; APPLICANT: Grund, Alan D.
; TITLE OF INVENTION: LINOLEATE ISOMERASE
; FILE REFERENCE: 3161-20-C1
; CURRENT APPLICATION NUMBER: US/09/561,077C
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/141,798
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 3684
; TYPE: DNA

```

```

; ORGANISM: Lactobacillus reuteri
; US-09-561-077C-25

```

```

Query Match 0.8%; Score 41.6; DB 3; Length 3684;
Best Local Similarity 50.0%; Pred. No. 0.68;
Matches 130; Conservative 0; Mismatches 129; Indels 1; Gaps 1;

```

```

Qy 4744 CAAATGGCGCCGAAAGAAAGTCAAGGCGCATCCCATCTAATGTGACGAGAAAACCTA 4803
Db 3342 CATGGTGGGAGATTTTAAAGCTATTGGATGACATCTACTAAAGCAGCAACATTA 3401
Qy 4804 TTTATACATAACCGGAACCTAAACCGGTAATCCGAAAGTACTTAAATTTAGCCA 4863
Db 3402 TTGAACTAGAGAAACAAAGTTAGAAAGAAATTAATTAATTAATTAATTAATTCGA 3461
Qy 4864 ACTCGCAATCTCACTTCTACATGAGTTGATGCTTTTTCACAGCTCTTAATTA 4923
Db 3462 GCGAAGAAATGACCTTACCGGAGCGGAATACCGCTTTGGCAAGCCAAATTAATTTTA 3521
Qy 4924 ACGCTGATTCCTGTTTAAATTTGTAAGCCTATATACCGCTCTTAAAGAACTCAGACGTGT 4983
Db 3522 TTAATTAATGTG-TAAACTTGACAGGCGCAGATACCTGTTTGAAGGGAGCTCAGAAAGCT 3580
Qy 4984 GCTTTTCTACTTTGTTTCA 5003
Db 3581 GCTTTTCTACTTTGTTTCA 3600

```

```

Search completed: November 30, 2005, 00:50:42
Job time : 594 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 14:49:57 ; Search time 211 Seconds
(without alignments)
3106.887 Million cell updates/sec

Title: US-10-030-850-2
Perfect score: 7611
Sequence: 1 MAYRKPSDLDGFICQMPKAD.....RNCISGGSGSTKNSAASSS 1492

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Genesegp21:*
2: genesegp1980s:*
3: genesegp1900s:*
4: genesegp2000s:*
5: genesegp2001s:*
6: genesegp2002s:*
7: genesegp2003as:*
8: genesegp2003bs:*
9: genesegp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7611	100.0	1492	4 AAB31597	Aab31597 Amino aci
2	1707	22.4	1526	3 AAB29634	Aab29634 Human pol
3	1352.5	17.8	1298	7 ADf60378	Adf60378 Human con
4	1348	17.7	1324	7 ADf60378	Adf60378 Rat putat
5	1348	17.7	1324	9 ADX07258	Adx07258 Cyclin-de
6	1344.5	17.7	1295	4 AAM39287	Aam39287 Human pol
7	1344.5	17.7	1309	4 AAM39286	Aam39286 Human pol
8	1344.5	17.7	1317	4 AAM41072	Aam41072 Human pol
9	1344.5	17.7	1317	4 AAM41073	Aam41073 Human pol
10	1214.5	16.0	916	7 ADB64143	ADB64143 Human pro
11	920	12.1	723	8 ADU20651	Adu20651 A. thalita
12	608.5	8.0	1330	8 ADU20651	Adu20651 Human pro
13	502.5	6.6	454	7 ADB63929	ADB63929 Human pro
14	497.5	6.5	432	5 ABB81781	Abb81781 Proteasom
15	328	4.3	1972	7 ADf68457	Adf68457 Human hea
16	328	4.3	1972	9 ADX06899	Adx06899 Cyclin-de
17	314.5	4.1	2038	6 AAE27901	Aae27901 Human LBD
18	304.5	4.0	2040	6 ABR83341	AbR83341 Human NOV
19	281.5	3.7	306	7 ADJ70616	AdJ70616 Human hea
20	276	3.6	407	8 ADY12968	AdY12968 Plant ful
21	232.5	3.1	2018	4 ABB71336	Abb71336 Drosophil
22	215	2.8	1684	2 ADH11336	Adh11336 Vertebrat
23	215	2.8	2340	2 ADH11419	Adh11419 Human hu
24	215	2.8	2365	7 ADD93763	Add93763 Human hel

25	215	2.8	2429	7 ADD93765	Add93765 Human hel
26	214	2.8	2429	3 ADY98003	AdY98003 Nav2 prot
27	210.5	2.8	2432	3 AAY85565	Aay85565 Human hom
28	205.5	2.7	1831	6 ABU43109	AbU43109 Protein e
29	205.5	2.7	1870	6 ABJ19019	AbJ19019 Pathogen
30	204.5	2.7	2344	4 AAU37120	Aau37120 Staphyloc
31	201.5	2.6	2261	6 ABJ18914	AbJ18914 Pathogen
32	201.5	2.6	2271	6 ABM72734	AbM72734 Staphyloc
33	196.5	2.6	2283	6 ABP56876	AbP56876 Staphyloc
34	196	2.6	199	8 ADU20428	AdU20428 A. thalita
35	195	2.6	2271	6 ABU16000	AbU16000 Protein e
36	194.5	2.6	2681	6 ABJ19025	AbJ19025 Pathogen
37	193.5	2.5	2137	5 ABP39618	AbP39618 Staphyloc
38	193.5	2.5	2137	8 ADS05656	AdS05656 Staphyloc
39	193.5	2.5	3201	4 ABB59129	Abb59129 Drosophil
40	192	2.5	1150	4 ABB59129	Abb59129 Drosophil
41	192	2.5	4274	4 ABG00972	AbG00972 Novel hum
42	192	2.5	4397	4 ABG21944	AbG21944 Novel hum
43	190.5	2.5	2426	2 ADH11285	Adh11285 Vertebrat
44	187	2.5	1079	7 ADJ71179	AdJ71179 Human hea
45	187	2.5	4386	4 ABG07375	AbG07375 Novel hum

ALIGNMENTS

RESULT 1.

AAB31597
ID AAB31597 strand1; protein; 1492 AA.

AC AAB31597;
XX

DT 30-APR-2001 (first entry)
XX

DE Amino acid sequence of a Drosophila polypeptide designated orbit.
XX

KW Orbit; chromosome segregation; mitotic spindle; microtubule; mitosis;
XX polyploid cell; microtubule binding; nucleation; motor protein; cancer;
XX leukaemia cell; solid tumour.
XX

OS Drosophila sp.
XX

EH Key Location/Qualifiers
XX

FT Domain /note= "HRI domain"
XX

FT Domain /note= "putative GTP binding domain"
XX

FT Domain /note= "putative GTP binding domain"
XX

FT Domain /note= "HRI domain"
XX

PN WO200104295-A1.
XX

PD 18-JAN-2001.
XX

PF 11-JUL-2000, 2000WO-GB002662.
XX

PR 13-JUL-1999, 99GB-00016402.
XX

PR 24-DEC-1999, 99GB-00030707.
XX

PA (UYDU-) UNIV DUNDEE.
XX

PI Avides MDC, Deak P, Glover DM;
XX

DR WPI, 2001-138345/14.
XX

DR N-PSDB; AAF24991.
XX

XX Novel Drosophila orbit protein and its human homolog, useful for treating
PT tumors and in assays for identifying substances capable of inhibiting
PT mitosis.

XX Claim 5; Fig 7; 92pp; English.

PS The present sequence represents *Drosophila* polypeptide designated orbit.
 CC The orbit protein is required for chromosome segregation. Orbit protein
 CC localises to mitotic spindles and binds microtubules. Orbit mutant
 CC phenotypes result in embryos exhibiting abnormal mitosis and polyploid
 CC cells. Orbit polypeptides are useful for inhibiting mitosis, and for
 CC identifying substances which affect orbit functions such as microtubule
 CC binding, microtubule organizing centre, nucleation activity and
 CC interactions with microtubule motor proteins. Tests for detecting or
 CC sequencing orbit, or its homologue, in a biological sample may be used to
 CC determine orbit sequences within cells in individuals who have or are
 CC suspected to have, an altered orbit gene sequence, for example within
 CC cancer cells including leukaemia cells and solid tumours such as breast,
 CC ovary, lung, colon etc

XX Sequence 1492 AA;

Query Match 100.0%; Score 7611; DB 4; Length 1492;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAYRRPSLDLDFIQMPKADRVKQVLAEDLVTFISDDTNSIVCTDMGFLIDGLMPALTG 60
 DB 1 MAYRRPSLDLDFIQMPKADRVKQVLAEDLVTFISDDTNSIVCTDMGFLIDGLMPALTG 60
 QY 61 SHFKIAQSLAEFSLIKRIGSDFNAYATVLPVYIDRLGSDRTVREKAQLRLDLMEH 120
 DB 61 SHFKIAQSLAEFSLIKRIGSDFNAYATVLPVYIDRLGSDRTVREKAQLRLDLMEH 120
 QY 121 RVLPFOALIDKLATSCFGRKNAKVEEFLQITVNAHEGYTQOLSVRVYIPVLCALGDP 180
 DB 121 RVLPFOALIDKLATSCFGRKNAKVEEFLQITVNAHEGYTQOLSVRVYIPVLCALGDP 180
 QY 121 RVLPFOALIDKLATSCFGRKNAKVEEFLQITVNAHEGYTQOLSVRVYIPVLCALGDP 180
 DB 121 RVLPFOALIDKLATSCFGRKNAKVEEFLQITVNAHEGYTQOLSVRVYIPVLCALGDP 180
 QY 181 TVNREAAIQTLVETVYKGVGRLRMDVPSKIAMLEOKDYOKOELLPSALK 240
 DB 181 TVNREAAIQTLVETVYKGVGRLRMDVPSKIAMLEOKDYOKOELLPSALK 240
 QY 241 NTNGNGVGLDEADNIGLERPPTRMKRPPLSAVSSSLRPKNVNDVTGADAVTMESPES 300
 DB 241 NTNGNGVGLDEADNIGLERPPTRMKRPPLSAVSSSLRPKNVNDVTGADAVTMESPES 300
 QY 301 SFEVVPQINIFHAKMDIYKQVLVIISDKNADWEKRVDALEKTRALLIISYHTOPFVA 360
 DB 301 SFEVVPQINIFHAKMDIYKQVLVIISDKNADWEKRVDALEKTRALLIISYHTOPFVA 360
 QY 361 VOLKELSLFVDILKEELRSQVIRACITIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
 DB 361 VOLKELSLFVDILKEELRSQVIRACITIAVMSKTLNKLDAFCWSLLEHLINLIONSA 420
 QY 421 KVIASASTIAKYIITKTHAPKLKIYDTLNQSKDIRSTLCELMWLFEEMQTKALE 480
 DB 421 KVIASASTIAKYIITKTHAPKLKIYDTLNQSKDIRSTLCELMWLFEEMQTKALE 480
 QY 481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
 DB 481 RNAIVLRDTLTKKSIQDADCDARRHSRYAYMAFRHHFPELADQIYGTLDIAAQRALERERE 540
 QY 541 GGGGGGTGTGTATPRTYRSIRGTGTLOKPTPSKRSISAVDTAAQAQAKRAQYTL 600
 DB 541 GGGGGGTGTGTATPRTYRSIRGTGTLOKPTPSKRSISAVDTAAQAQAKRAQYTL 600
 QY 601 YSRQKPLGPNNSNOASMTGAASGLPRPLNSNGGTPATTPSVPVPRGRGAGVSQS 660
 DB 601 YSRQKPLGPNNSNOASMTGAASGLPRPLNSNGGTPATTPSVPVPRGRGAGVSQS 660
 QY 661 QPGSRSTPSTKLADQYGIQNYRGATGAI PKKASGIPRSTASSRETSPTRSGGGLMKR 720
 DB 661 QPGSRSTPSTKLADQYGIQNYRGATGAI PKKASGIPRSTASSRETSPTRSGGGLMKR 720
 QY 721 SMYSTGASRRTPEPNNVPRPSAPKLLAQSREAHITLGVDDQGPVYSGDYMRSGGMR 780
 DB 721 SMYSTGASRRTPEPNNVPRPSAPKLLAQSREAHITLGVDDQGPVYSGDYMRSGGMR 780

DB 721 SMYSTGASRRTPEPNNVPRPSAPKLLAQSREAHITLGVDDQGPVYSGDYMRSGGMR 780
 QY 781 MGRKLMGRDSDIDSEASVCSERSPDSSYTRGNKSNYSLSGSHRLDMSTORAPDDI 840
 DB 781 MGRKLMGRDSDIDSEASVCSERSPDSSYTRGNKSNYSLSGSHRLDMSTORAPDDI 840
 QY 841 ETTIQFCASHTWSEKKGILSTLOYLADGKELTQOOLKCVLDNRKPMQTHTVYSLFL 900
 DB 841 ETTIQFCASHTWSEKKGILSTLOYLADGKELTQOOLKCVLDNRKPMQTHTVYSLFL 900
 QY 901 DTVLELVHANESRNGSSCLTRLPNKLGTDILNMSKIMKTLDVHEYPPTOLQK 960
 DB 901 DTVLELVHANESRNGSSCLTRLPNKLGTDILNMSKIMKTLDVHEYPPTOLQK 960
 QY 961 ELFRISDSTQPTTKTRIALRLFTDLANTYCKSSDPSPDOSQACERTYKLAQLAADQ 1020
 DB 961 ELFRISDSTQPTTKTRIALRLFTDLANTYCKSSDPSPDOSQACERTYKLAQLAADQ 1020
 QY 1021 KSMELRSQARSCVVALYNLMTPOWTLILLADLPKYQDSASCIHSNRRSQSCNSGANS 1080
 DB 1021 KSMELRSQARSCVVALYNLMTPOWTLILLADLPKYQDSASCIHSNRRSQSCNSGANS 1080
 QY 1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISTSPSRSSVBOELLFSELDIOH 1140
 DB 1081 PSSSPSSSPKPLQSPSVGPASLOSHHQLSISTSPSRSSVBOELLFSELDIOH 1140
 QY 1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOODSCASLSNSKTQSSANTQSN 1200
 DB 1141 NIQTSSEIRHCFQGYOTALAPNGFNGLQYHDQGOODSCASLSNSKTQSSANTQSN 1200
 QY 1201 TPESATWRLNLEBERTTQNAKSPTDKATITVYSINAEENGELILASNLMESEVVAALT 1260
 DB 1201 TPESATWRLNLEBERTTQNAKSPTDKATITVYSINAEENGELILASNLMESEVVAALT 1260
 QY 1261 LTKDQPELLOSLTNIGICIKGNCCLPKHFRSIRMLNIIIEAHTDVVLAGLHVS 1320
 DB 1261 LTKDQPELLOSLTNIGICIKGNCCLPKHFRSIRMLNIIIEAHTDVVLAGLHVS 1320
 QY 1321 KIMSNNKRRNMHFLLELLKIIQCYOHSKEALRDISMIPIRAPSLPLDSINIVNPV 1380
 DB 1321 KIMSNNKRRNMHFLLELLKIIQCYOHSKEALRDISMIPIRAPSLPLDSINIVNPV 1380
 QY 1381 IATGEPTNLCATKILLETVEHSGEITDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
 DB 1381 IATGEPTNLCATKILLETVEHSGEITDAHLDIVPNNLARSADDTQSMVRKAAPFCIVK 1440
 QY 1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492
 DB 1441 LYFVLGEKVKPKLSVLNPSKVRLLNVIIEKORNCISGGSSSTKNSAASS 1492

RESULT 2
 AAB29634
 ID AAB29634 standard; protein; 1526 AA.
 XX AAB29634;
 AC 21-FEB-2001 (first entry)
 DT
 XX
 DE Human polliosis-associated gene 373-encodded protein, SRQ ID NO:2.
 XX
 KM Human; polliosis-associated gene 373; IgF; immunoglobulin E;
 KM cedar pollen allergy; T-cell; reduced expression; detection; diagnosis;
 KM drug screening; allergic disease.
 OS Homo sapiens.
 XX
 PN W0200065046-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 26-APR-2000; 2000MO-JR002730.
 XX

PR 27-APR-1999; 99JP-00120489.
 (GENO-) GENOX RES INC.
 PI Nagasu T, Sugita Y, Kaehiwabara T, Oshida T, Obayashi M, Gunji S;
 PI Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K;
 XX WPI: 2000-687339/67.
 DR N-PSDB; AAC64201.
 PT Pollinosis-associated gene 373 undergoing significantly low expression in
 PT subjects with high cedar pollen-specific immunoglobulin-E levels, useful
 PT in diagnosis of allergic diseases and screening drug candidates.
 PS Claim 1; Page 58-69; 80pp; Japanese.
 CC The invention relates to the human pollinosis-associated gene 373 which
 CC exhibits significantly reduced expression in the T-cells of individuals
 CC with high cedar pollen-specific IGE (immunoglobulin E) levels. The gene
 CC was isolated from T-cells from individuals allergic to cedar pollen using
 CC the differential display method. The invention also relates also relates
 CC to the protein encoded by pollinosis gene 373; expression constructs and
 CC host cells comprising pollinosis-associated gene 373 nucleic acids;
 CC pollinosis-associated gene 373 primers and probes; antibodies against the
 CC protein encoded by 373 nucleic acids; and a method of diagnosis of allergic
 CC diseases via the detection of pollinosis-associated gene 373 nucleic
 CC acids. The invention additionally encompasses methods of screening drug
 CC candidates for the treatment of allergic disease by measuring the
 CC expression of pollinosis-associated gene 373 in pollen antigen-stimulated
 CC T-cells in the presence of a test compound relative to a control.
 CC Pollinosis-associated gene 373 is useful in the diagnosis of allergic
 CC diseases and in the screening of drug candidates for the treatment of
 CC such diseases. The present sequence represents a protein encoded by human
 CC pollinosis-associated gene 373
 XX
 SO Sequence 1526 AA;
 Query Match 22.4%; Score 1707; DB 3; Length 1526;
 Best Local Similarity 30.1%; Pred. No. 6,3e-119; Indels 206; Gaps 45;
 Matches 485; Conservative 300; Mismatches 618;
 QY 5 KPSSDGLIQMPKDMKRVQLAEDLVFL-SDDTNSIVCTDGMFL--IDGLMPWLTG 60
 DB 1 EPRSEHYCAOYQKQDVGRLVGOBLLYLGPALISDEEDLRLGKTVADALGWGVS 60
 QY 61 SHFKIAQSLKFAFSELIKRLGSDFNAYTATVLPVHYIDRLGSRDLYREKAOILLRLMEH 120
 DB 61 SNYKVSIMQLEILSAFVDRILSTRFSKYVMIVIALIDRMGDADKDKRDEAOILLKMD- 119
 QY 121 RVLPPQALIDKLATSCFKRKAQKVRREFIQTIVNLHEGYQQLSVRVYIPVLCALGDP 180
 DB 120 QVAPRYIWEOLA-SQFKKRNFRSREGVCLCLTINIGAOPLVYSKILPHILCLIFGDS 178
 QY 181 TVNVEEAQOTLVEIKVGDRLRPLRMDVPAKSLMLQKEDQVQEGILLPSALK 240
 DB 179 NSQVRDAALAIYEIRAHGKRMOLYKR-GIPPARLEWIFAKFDEVQSSGGMILSVCK 237
 QY 241 NTNGNGVGLDEADNIGLRERPT-----RMIKRPLHSAVSSSLRPPKPVNDVT-- 287
 DB 238 DKS-----FDDDESVD-GNRPSAASAFKVPAPKTCGNPANSARKPGSAGPVGASAE 291
 QY 288 GDGAGVATMESSESSFEVYPOALNFHAKMDMDIKOVUVIISDKNADMEKRVDAIKKIRAL 347
 DB 292 GGAGVADDEDFIFAFDVPISIQYSSREELTNKRIEILSDKHDQDQANLKKIRSL 351
 QY 348 LI---LSYHTQGFVAVQVQLKSLSPVDIKELRSQVIREACTITIAVSKTLKTLADAF 404
 DB 352 LVGAAQYDGFQHLRLLDALKLS-----AKDLRSQVREACTIVAHLSLVGNKFDHG 406
 QY 405 CWSLIEHLINLIONSASAVIASASTIALKYIKYTHAPKLKITYDTLNOGSKSDIRSLC 464
 DB 407 AEAIIVPLFPLVPSAKVMATSCGAIRFIIRHTHVRLPLILITSNCT-SKSVFVRRRSF 465

QY 465 ELMTVLLFEEMOTKALERNATVLRDLTKKSIGDADCDARRHRRYAVMAFRHPPELADQY 524
 DB 466 EFLDILLDEMGTHSLERNAVALVETIKKGIHDADAEARVEARKTYGRLNHPGEAETLY 525
 QY 525 GTLIDIAQPALEREREGGG-----GGTGTGTGTAPETRR 559
 DB 526 NSLEPSYQKSLQTLTKSSGSVALPQSDRSSSSQESLNPRESSKMWSTANPESTVAGRVA 585
 QY 586 TVSRIGRPTGLQKTPMSRSISAVDTAAQAKRYAOTVLYSRQKPLGPNNSNOAS-- 617
 DB 586 GSKKASLPQSLQ-----KRSNIDVNAAGAKAHHAAGSVRSGR-LGAGALNAGSYA 638
 QY 618 -----MTGAAGSLPRPLNSNGTPTATPGSVTPPRRAG--VQSQPGSS 666
 DB 639 SLEDTSDKLDGTASDEGRVAKLSA-----PLAGMNAKADRGSRRTKRVQSQPGSS 693
 QY 667 TSP-----STKLRDQYIGNYRGATGALPKKASGIPSTIASRPSPTR--SGGGLM 718
 DB 694 GSPGRVLTTLTALSTVSSGVQVLYVNSASA--OKSKIPRSQCSREASPSRLSVASSRI 751
 QY 719 KSMSTGA-----GSRTPRRNPNVPSAP-----A 745
 DB 752 PRPSVQSGSREASRESRSDTSPNSFQPLASHHSRTGALYAPEVYAGSGGYGISQ 811
 QY 746 RLIAQREAEHTLVGDDQPDVSGDYRSGGMRGKLMGEDESDDIDSEASSVCSER 805
 DB 812 SRLSSVSAMRVLTNTSD--VEEAVADALKKPARRR-YESYGMHSDDANSDASSACSER 868
 QY 806 SPDSYATRNKSNVSLSGSHTRLDNSTORAPPDITIIIOFCASTHSEKDGILSTQY 865
 DB 869 SYSS-----RNGSI-----PYMKGTDVAEVLNRKSSNMSEKGLIGLQNL 912
 QY 866 LAQKELTQOQLCVLDMFERKFMEDHTK--VSLFLDTVTELLVHANETSRNGSSCLT 924
 DB 913 LKQRTLSRVELKRLCGIFRMRADPHGKVPFMSFLETLVDPIQVHDDL-QDWLFVILT 971
 QY 925 RLFWLGTDLNLSHNSKIMWTLQVNHETPQOLKELFRIISDSTQPTTKRIALIRF 984
 DB 972 QLKMKADLLGSQAVQKALQVDTRESFPNDQFNIMFTVQDTQTPSLKVALILKY 1031
 QY 985 LTLDLANTYCKSDPSD--SQACERVYLKLAOLAQKMELEASQASCVLALYNLPT 1042
 DB 1032 IETLA-----KOMD-PGDFINSEETRLAVSVITYTTTPKSDVKAQSVLISPELNT 1086
 QY 1043 QMTLLADLPVNYDQASRCSHSMRROSQSCNAGSPSSPLSSSPK-----PLQ 1095
 DB 1087 EFTMLGLPPTFDGATKLLHNHR--NTNGTQSSMGSPLRPTRRPAMWSPLT 1142
 QY 1096 SPASGPRFASLQSHHQLSISSTSPRSQSV-----QELLFSELDIQHNIOKTSSE- 1148
 DB 1143 SPYNTSONTLPSAPVDVTEMNSEDIYSLRGVTEALIONFPSQOEDMNEPLKDSKDD 1202
 QY 1149 --IRHCFGQYQYALAPNGFNGLQYHDQGOQDSCASLSNKSQSSANTQSN----TP 1202
 DB 1203 DGDMSCG--PGKSDPRAGDADTDSQOTALDNKASLSHMPHSSPRSRDYNPNYNSD 1258
 QY 1203 ESATKRLNLERERTQOAKSPTDQAKVITYSINNAENGELILASINSESEVVRVALTLT 1262
 DB 1259 SISPNKSLKXAMDADDADQPPD-----LSLSDS--LVAELEKLS-----N 1301
 QY 1263 KDQPELLOTS/TLNGICIKGNCELPNKHFRSIRMMLNLLEAHTDVVINGLHLSKI 1322
 DB 1302 KHERVEERKIALYELMLKQESPSVMBEHTLILLLETLGDEKPTFALALKVLREI 1361
 QY 1323 MRSNKRNMHMFLEILIKITQYQH-SKELRQIDSMIPRIAPSLPDLGINVNPY 1381
 DB 1362 LKHQPAR--FKVYAEITVWKLEAKDPRKEVVAEAAEASVLAISISPECICVLCPII 1419
 QY 1382 ATGEPTNLCALKILEVEHNGSEITDAHLDIVFPNLARSADDTQSMVKAAVFCTVCL 1441
 DB 1420 QTAADYDINLAIMQTVIERVSKETLNLLEBIMBGLQGYDNSSSVKXCVFCLAV 1479

QY 1442 YFVLGEKVKPKLSVNPKNVLLNYTEKORNCISGGSSTKNSSAS 1490
 DB 1480 HAVIGDE-LKPHLSQLTGSKMLNLYI-KRAQTGSGADPTDVGOS 1526

RESULT 3
 ADF60378
 ID ADF60378 standard; protein; 1298 AA.
 XX ADF60378;
 AC ADF60378;
 DT 12-FEB-2004 (first entry)
 XX
 DE Human contig polypeptide sequence SEQ ID NO:2745.
 XX
 KM biological activity; genetic engineering; hybridisation probe; oligomer;
 KM primer; chromosome mapping; gene mapping; recombinant protein production;
 XX human.
 OS Homo sapiens.
 XX
 PN WO2003080795-A2.
 XX
 PD 02-OCT-2003.
 XX
 PF 09-AUG-2002; 2002WO-US025485.
 XX
 PR 09-AUG-2001; 2001US-0311261P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Yang Y, Wang Z, Wang G, Ma Y;
 XX
 DR WPI; 2003-876918/81.
 XX
 DR N-PsDB; ADF59926.
 XX
 PT New polynucleotides, useful as hybridization probes, oligomers or
 PT primers, for chromosome or gene mapping, for the recombinant production
 PT of proteins, and for generating antisense DNA or RNA.
 XX
 PS Example 3; SEQ ID NO 2745; 571pp; English.
 XX
 CC The present invention describes isolated polynucleotide sequences (1),
 CC which encode polypeptides (11) with biological activity. Also described:
 CC (1) a vector comprising (1); (2) an expression vector comprising (1); (3)
 CC a host cell genetically engineered to comprise (1) which is operatively
 CC associated with a regulatory sequence that modulates expression of (1) in
 CC the host cell; (4) a polypeptide (11) encoded by (1); (5) a composition
 CC comprising the polypeptide of (4) and a carrier; (6) an antibody directed
 CC against the polypeptide of (4); (7) detecting (1) or the polypeptide of
 CC (4) in a sample; (8) identifying a compound that binds to the polypeptide
 CC of (4); (9) producing the polypeptide of (4); and (10) a collection of
 CC polynucleotides comprising at least one of the polynucleotide sequences
 CC (1). The polynucleotides (1) can be used as hybridisation probes,
 CC oligomers or primers, for chromosome or gene mapping, for the recombinant
 CC production of proteins, and for generating antisense DNA or RNA. The
 CC present sequence represents a human contig polypeptide sequence, which is
 CC used in an example from the present invention.
 XX
 XX
 XX
 SQ Sequence 1298 AA;

Query Match 17.8%; Score 1352.5; DB 7; Length 1298;
 Best Local Similarity 29.9%; Pred. No. 3.1e-92;
 Matches 415; Conservative 235; Mismatches 496; Indels 243; Gaps 47;

QY 236 PSALNTNGNGVGLADNIGLRERPTNKRPLHSAYSSSLRPKNNDVTDGAGATM 295
 DB 15 PSSRN-----VGGTTRRLG-----SLLGSKSA--AKEGGAUDE 50
 QY 236 ESSESSFEVVPOLNIFHAKMDIYKOVLIISDNKADMEKEVDLKKIRALLIISYHQ 355
 DB 51 EDFIAFDVDPVVOIYSSRDLESINKIREILSDKDHMEQKRVNMLKKIRSLLAGAAFY 110

QY 356 POFAVQLKEISLSTVDILKEELRSQVIREACTITAWSKTLRNKLDAPFCSILEHLNL 415
 DB 111 DNFFQ-HILRLDGA-F-KLSAKDLRSQVVRACITLGHLSVYLGNFHDGAELMPTIFNL 168
 QY 416 IONSAKIVASASTALKTIYKYTHAPKLIKITYDTLNSKSKDRISTICEMLVLLPEEQ 475
 DB 169 IPIISAKIIVATSGVAVARLIRHTHPLIPVITSNCT-SKAVAVRRCPPEFLDLLOEQ 227
 QY 476 TKALERNATVLRDLTKKSIQDADCDARRSHRYAWAFRRHPELADQIYGTLDIAAQAFL 535
 DB 228 THSLERHISVLAETIKKGIHDADSEARIEARKCYWGFHSHFSRAEHLVHTLESSYQAL 287
 QY 536 ERENKGCGGGGTGTCTGTAPETRTVSRIGTPTGLQKPTSMGISAVDTAAQRAKYR 595
 DB 288 QSHLK-----NSDIYSLPQSDRSS--SSQSLNRLPLSAKSPSTGTTISRASTVSTK 338
 QY 596 AQYLTYSRQRKPLPNNNSQAMTGAAGSLPRPLNSNGGTP-----ATTPEGV--- 647
 DB 339 SVSTGSLQR-----SRSDIDVNAASA--KSKVSSSGTTPSSAALLPQSVASL 388
 QY 648 -----TPRPRAG-----VSQSQ-----PGSRSTSP 669
 DB 389 DGTITKAGRIITRRQSSGATNVASTPDNNGRBRKAVVQSQSRSANPAGASRSSSP 448
 QY 670 STKLRDQYGI-GNYTRG-ATGAI PKKASGIPRSTASREFTSPTRSG--GGLMKRSNY 724
 DB 449 GKLISGCGYGLTGSSSRGPVTPSEKRSKIPRQSGSRETSFNRIGLARSRIIPRPSMS 508
 QY 725 TGAG-----SRTPENNNPRPSAPARLLAQSRFAHTIGVGDGQ-PDVSDDYRSGG-- 778
 DB 509 QCGSRDTSRESSRDTSPARKFPPL-----DRFLGQPGRI PGSVNMRVLSSTD 558
 QY 779 -----MEMGRKLMGRDESDIDISEASSVCSESPSSYTRGNKSNYLSGSH 826
 DB 559 LEAVADALKKPRVRRRPEYMTSDDANGASSVCERSYGS--RNGCIHY----- 609
 QY 827 RLDMSTQAPRPDIETIIQFCASTHSEBKDGLISLQYLDAGKELTQOOLKCYLDNFRK 886
 DB 610 -----LRQTEDEVAEVLNHCASSMSERKEGLGLGNLMSQRTLSVEHLKRIICEIFTR 662
 QY 887 MEMDTHK-VYSLFLDPVTTELILVHANEITSNGSSCITRFNLG--TDLNSMSKRW 944
 DB 663 MFADPHSRVFSMLETIVDPIIHKODL-ODWLFVLTQLKXGAEADLGSVQAKQK 721
 QY 945 TLQVVEHYFTQLQKELFRILISDSTQPTTKTRIALRPLTDANTYCKSSDPSD--Q 1002
 DB 722 ALDVTROSFPDQFNILMRITVDQGTQPNLKVVALIKYTESLA---RQMD-PTDEVN 776
 QY 1003 SQACERYVLKLAQADQKSMELRSQARSCIVALLYNTPMTLLDLPVYQDSARSC 1062
 DB 777 SSETRLAVSRITITTEPKSDVAKAQIVLISFEILNTEPFTMLGALPTFPDGAIVKL 836
 QY 1063 IHSIMRQSQSCNSGANSPPS-----SPSSSGPKLQSPS-----VGFFA----- 1103
 DB 837 LNLHLKMS--NVSQSVSPWTIGTRPSRHTSSTPSTSTNCSHGSLSRIMGMSAD 893
 QY 1104 SLQSHHQLSISSTSPRSOSVYQELLFSSELDIQ-HNIQKTESEIRHCGGYQTALA 1162
 DB 894 GLAHPPFPQSPNIPFAPBHKALRYSRPSMDYDENT--NSEELYSLSRGTE-AIE 950
 QY 1163 PNGFNHLYHQDQGDQ--SCASLSNSKTSQSANQTQNTP-ESATMRIDNLERERT 1218
 DB 951 KFSFRSQEDNEPIKRGKKCEDIVSDGGAASAPATGRGSGSEVEGRTALDNKTSILNT 1010
 QY 1219 QNAKS-----PTDDAKITVISIMANGEL-----IL 1245
 DB 1011 QPPRAPFPAPRDYNPFPYSDA-----INTYDTALKEAVFPDDMEQLRDVPIDHSDVL 1064
 QY 1246 ASNIMSESVRVALLTKQDPVELLO-TSLTNLIGICIKGNCELPNGHFRSIRMMLNITL 1304
 DB 1065 ADLKEISNNEREERKGAALLBELIKITREDSDGVM-----BEHFTIILLLETL 1115
 QY 1305 EAEHTDVVIALGHLVLSKIRSNQRHNMWFLLELILIKTIQCYOHS-KEALRDIDSMIPR 1363

```

Db      1116 GDSKHSIRALATKLVLEILNQPAP--FKNYAELTITMKTLEAKHDSHKEVWPAEPAAST 1173
Qy      1364 IASLPLDLISINIVNPIATGEPTNLCATKILETEHHGSEITDAHLIVPPLARSA 1423
Db      1174 LASSIHEOCITKVLCPITOTADPFINLAIMKOTKVERAKSLLOLVDIIPGLLOQ 1233
Qy      1424 DPTQSWRKAAFCIVLYFLGEEKVPKLTVLNPBKVLLNVYIEKORNCISGGGSS 1483
Db      1234 DNESSVRSKASVFCIVAIYVIGED-LKPHLAQLTQSKMKLNLNLYIKR-----AQT 1283
Qy      1484 KNSSAASS 1492
Db      1284 TNSNSSSS 1292

RESULT 4
ADB79873 standard; protein; 1324 AA.
XX
XX ADB79873;
XX
XX 04-DEC-2003 (first entry)
XX
XX Rat putative KIA0627, SEQ ID 113.
XX
XX Analgesic; pain; streptozocin-induced diabetes; rat.
XX
XX Rattus norvegicus.
XX
XX EPI279744-A2.
XX
XX 29-JAN-2003.
XX
XX 26-JUL-2002; 2002EP-00255249.
XX
XX 27-JUL-2001; 2001GB-00018354.
XX
XX 07-FEB-2002; 2002GB-00002910.
XX
XX (WARN ) WARNER LAMBERT CO.
XX
XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
XX
XX MPI: 2003-395407/38.
XX
XX N-PSDB; ADB79874.
XX
XX Use of isolated gene sequences and encoded polypeptides that are
XX upregulated in the spinal cord in response to streptozocin-induced
XX diabetes for screening compounds for the treatment of pain, or for
XX diagnosing pain.
XX
XX Claim 1; Page 204-208; 334pp; English.
XX
XX The present invention relates to nucleotide sequences which are useful in
XX the screening of compounds for the treatment of pain, or for the
XX diagnosis of pain. The nucleotide sequences are up-regulated in the
XX spinal cord in response to streptozocin-induced diabetes. The present
XX sequence was used to illustrate the invention.
XX
XX Sequence 1324 AA;
XX
XX Query Match 17.7%; Score 1348; DB 7; Length 1324;
XX Best Local Similarity 29.6%; Pred No. 6.9e-92;
XX Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;
Qy      237 SALK-----NTNGNGVGLDEADNIGLERPTRMIKRPLHSVSSSLRPKNVNDVTDGAGA 292
Db      55 SAFKVPAPKTSGN-----PANSARKP-----GSAGGPKVGASKEGAGA 94
Qy      293 VTMESEFSEFVVPQUNITHAKDMMDIYQVLIISDKNADWEKRYDALKTRALLI--- 349
Db      95 VVEDDFIKAFIDVPSIOIYSSRELBTLNKIREIISDDKHWDQDANALKTRSLVAGA 154

```

```

Qy      350 LSYHTQPOFVAVOLKELSLSFVDLIKELRSQVIREACTITTAVMKTLRNKLDACWSTL 409
Db      155 AOIDCFQHTRLDLGALKLS-----AKOLRSQVREACTIVAHSLTVGNKEDHGAEALV 209
Qy      410 EHLNLIQNSAKVIASASTALKTIITKTHAPKLKITDTLINOKSKDIRETCLGMLV 469
Db      210 PTLFNLVPSAKVMAVTSQGAIRFIIRHIVRLPLITSNCT-SKSVPRRSSEFLLD 268
Qy      470 LREEMOTXALEANAVLTDLTKKSIQDADCDARRSRVYVMAFRBHFPELADQIYGLDI 529
Db      269 LIQEWQTHSLERHAALVETIKKGIHDADAEARVARKTYMGLRNHFPEAEFLVNSLP 328
Qy      530 AAORALERELEGCG-----GGTGTGTGATPRTTRVSI 564
Db      329 SYOKSLQTYLKSSGSVASLPSQDRSSSSQESLNRPFSKMSSTANSTVAGVSAGS 388
Qy      565 GRTPGTLOKPTPSMSISAVDTAAORAKVRAQYTLYSQKRPLEPNNQMS----- 617
Db      389 SSLPGSLQ-----NSRSDIDVMAAAGAKH-HAAGQSVRGRGLGALNAGSASLEDT 441
Qy      618 ---MTGAALSGLPRRLNLSNGGTPATTGSAVTPRPRGAG---VSOGQPSGRSTSP-- 669
Db      442 SDKLDGTASEDEGRVRAKLSA-----PLAGMNAKADSRGRSRTKMYSQGPGSRSSG 496
Qy      670 ---STKLADQYGGIGNYVRGATGATPKASGIPRSTASRETSPTR---SGGLMKRSMY 723
Db      497 VLTTLTSLVSSGVQVLYNSASA--QKRSKI PRSGCSREASPSRLSVARSRIIPSV 554
Qy      724 STGA---GSRRTPERNNPVRRSAP-----ARLQA 750
Db      555 SQGCSREAGRESRDTSPVRSFQPLASRHSRSTGALYAEVYGAGPGYGISQSSRLSS 614
Qy      751 SREARTLVGDDGQDQDYSGDYMRGMRGKRLGMDSDPIDSEASVCSERSPDS 810
Db      615 SVSAMRVLTGSD--VEEAVADALKKPARR-YESIGMHSDDANDSASACERSYSS- 670
Qy      811 YTRGNSNYLSGSHTRLDWSTORAFDDIETIIQFCASTHWSEKDKGLSLTOYLADK 870
Db      671 -----RNGSI-----PYMRQTEDEVAEVLNRCASSNBSRKGLGLQNLNQR 715
Qy      871 ELTQOOLCVLDMFRKQFMDTHK-VYSLFLDVTTELIVHNETSNGSSCLTRLENK 929
Db      716 TLSRVELKRLCEIFTRFADPHGKRVSMLETLVDFIQVHNDL-ODWL FVLTOLDK 774
Qy      930 LGTDLNSMSKTKMTQVYHVEYFPQLOKELPRIISDSQPTTTRIALIRFLDIA 989
Db      775 MGADLLGSVQAKVQKALDVRESFPNDLFNILRFTVDOTOTPSLKVAIILKIETLA 834
Qy      990 NTYKSSDPFSD--SQACERTVLKLAQLADOKSMELRSQARSCLVALYNLTPQMTL 1047
Db      835 ---KQMD-PGDFINSSETRLANSRVITWTTERKSSDVRAKQSVLISLFLNTPPEFTML 889
Qy      1048 LADLPKYODSARSCLSHNRKROSQSCNSGNSPSSPLSSSRK-----PIQSPVIG 1100
Db      890 LGLAPKTFQOGATGLLNHLR-----NTNGNGOSMGSLPTPTPRSPAMWSPLTSPNT 945
Qy      1101 PFASLOSHHOLLSISSTSPRSQSSVE-----OELLFSSGLDIOHNOKTSEE---IRH 1151
Db      946 SQNTLSPSAPDYDTENNMSEDIYSLRGVTEALQNFSTRSQEDMNEPLKRSKDDGDSM 1005
Qy      1152 CFGGOYOTALAPNGFNGLQYHNDQODSCASLSNSKTQSANTQSN---TPESATW 1207
Db      1006 CGG---PGMSDPPRAGDADTSSQTALDNKASLSLHSMPTSSPSRDNPNVYSDISPF 1061
Qy      1208 RLNDLERRKTONAKSPTDRAKVITVSINNAENGELLIASLMSSEVVRVALTLTKDQPV 1267
Db      1062 NKSALEKAMPDDADQFDD-----LSLDHSD---LVAEILKELS-----NHNRV 1104
Qy      1268 ELIQTSLNLGICIKGNGCELPNKGFPSIMMILNIEAEHTDVYIAGIHLYSKIMRSNK 1327
Db      1105 EERKIALYEMLKLTQOESFYSWDEHFKTLLLETJGDKPPTRALALKVLEIRLQRP 1164
Qy      1328 MRHNMHLELILKTIQCYOH-SKEALRDIQSMI PRIAPSLPLDLSINIVNPIATGEF 1386

```

Db 1165 AR--FKYVAELTVNKTLEAHDPKHEVVRSAEASVATLSISPEOCIKVLCPIIQADY 1222
 QY 1387 PYNCAIKILEVEVHSEITDAHLDIVFNLARSADDDTOSWVRKXAVFCIVLYVLG 1446
 Db 1223 PINLAIKMQRKVERSKETLNLPEINBGLIGDINSESSVRKACVFLVAVHVIIG 1282
 QY 1447 EEKYKPKLSVLNPSKRVLLNVIYIEKORNCISGGGSSSTKNSAAS 1490
 Db 1283 DE-LKPHLSQLTGSKMKLNLNYI-KRAQTGSGGADPTTVDSGQS 1324

RESULT 5
 ADX07258
 ID ADX07258 standard; protein; 1324 AA.
 XX
 AC ADX07258;
 DT 21-APR-2005 (first entry)
 DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 1823.
 XX
 KM cytoeatic; cyclin-dependent kinase; cdk; biomarker.
 XX
 OS Homo sapiens.
 XX
 PN MO2005012875-A2.
 XX
 PD 10-FEB-2005.
 XX
 PF 29-JUL-2004; 2004MO-US024424.
 XX
 PR 29-JUL-2003; 2003US-0490890P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
 XX
 DR WPI; 2005-163068/17.
 XX
 DR N-PSDB; ADX07257.
 PT Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.
 XX
 PS Claim 5; SEQ ID NO 1823; 141pp; English.

This invention describes a novel method of predicting or determining whether a mammal will respond or is responding to an anti-cancer agent that modulates cyclin-dependent kinase (cdk) activity. The method comprises measuring the level of one or more biomarkers selected from 2774 biomarkers given in the specification (nucleotide sequence SEQ ID NO:1246 (Genbank EST W28729) is especially preferred). The method of the invention is utilized in a kit for determining or predicting whether a patient would be susceptible or resistant to treatment by an agent modulating cdk activity. The invention also describes a method for utilizing individualized genetic profiles for treating diseases and disorders based on patient's response and molecular level, specialized microarrays comprising the biomarkers described, antibodies directed against the biomarkers and a cell culture model to identify biomarkers. The cdk modulator is preferably N-5-[5-(1,1-dimethylethyl)-2-oxazolyl]methylthio-2-thiazolyl-4-piperidine carboxamide, 0.5-L-tartaric acid salt. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences. This sequence represents a biomarker used in the method of the invention.

Sequence 1324 AA;
 Query Match 17.7%; Score 1348; DB 9; Length 1324;
 Best Local Similarity 29.6%; Pred. No. 6.9e-97;
 Matches 404; Conservative 233; Mismatches 523; Indels 204; Gaps 39;

QY 237 SALK---NTNGVGGLDEADNIGLEREPTRMIKRPLHSAVSSSLRPKNVNDVTDAGA 292
 Db 55 SAKVPAPRTSGN-----PANSARKP-----GAGGPPKVGAGKEGAGA 94
 QY 293 YIMSEPFSEFVVPOLNIFHAKOMDDIYKQVLIISDXNADMKRVDALKKIATALLI--- 349
 Db 95 VDEDDFTKAFPTDVPSTQIYSSRELEETLNKIRIELISDDHDQORNALKKRLVAGA 154
 QY 350 LSYHTQPOFVAVOLKELSLFVDILKEELRSQVIREACTITIAVMSKTLRNLKLDAPCWSIL 409
 Db 155 AQYDCFPQHLRLDGLKXS-----AKDLRSQVIREACTITIAHLSVNLGNKPFHGAEATV 209
 QY 410 EHLINLIONSAXVIASSTIALKYIIKYTHAPKLYITDTONQSKDIRSTLCELMVL 469
 Db 210 PTLFNLVPSAKVMATSGCAIRFIIRHIVHPLRPLITSNCT-SKSVPRRRSPFELDL 268
 QY 470 LPEENQTKALERNATYLRDTLKKISIGDADCDARHRHRYVMAFRHFPPELADDIYGTLDI 529
 Db 269 LQEMQTHSLERHAVALVETIKKGIDHDAEAREARKTYMGLRNHFPGBAETLYNSLEP 328
 QY 530 AAOBALEREREGGG-----GGTGTGTGTAPETRRTVSRI 564
 Db 329 SYOKSLQTYLKSGSVASLPOSDRSSSSQESLNRPFSKMSYANSTYAGRVASAGSSKA 368
 QY 565 GRTPTGLQKPTPSMRSISAVDTAAAGRAVRAQYTLYSRQKPLGNNSNQAS----- 617
 Db 389 SSLPGLQ-----RERSDIDVNAAGAKAH-HAAGQSVRRGRGLGALNAGSVASLEDT 441
 QY 618 ---MTGAASGSLPRRLNLSNGGTPATTPGVSVPFRGRAG---VSOGQPGSRNSP-- 669
 Db 442 SDKLDGTASEDRVRKXLSA-----PLAGMGNAKADRSRSRTKVVSSQSPGSRSSPCR 496
 QY 670 ---STKLRDQYGAIGNYRGATGATPKKASGIPRSTASRETSPTR---SGGLMRSMY 723
 Db 497 VLTITLSTVSSGVQVRVLYNSASA--QKRSKIPRSGCGRERASPSLSVARSSRIIRPEV 554
 QY 724 STGA---GSRTPPERNNPVPSAP-----ARLLAQ 750
 Db 555 SGCSSREASRESRSDTSFVRSFOPLASRHSRSTGALVAPEVYASGPGYGISQSRLSS 614
 QY 751 SREAHTLGVDGDDGDDYSGDYMSGGMKRMKLMGRBESDIDSEASSVCSERFPDS 810
 Db 615 SYSAKRVLTGSD--VEEAVALKKPARR--YESYQMSDDNADASACSERYS- 670
 QY 811 YTRGNKSNYLSGSHRTLDWSTORAFPDDETIIQFCASTHMSERKDGHSITQVYADGK 870
 Db 671 -----RNGSI-----PTYMQTEDEVAVLARCASSNMSERKEGLGLGNLKNQR 715
 QY 871 ELTQQLKCVLDMFRKMPDHTK-VSLPDTVTLELIVHANETSRNGSSCLTRLPNK 929
 Db 716 TLSRVKLKLCIEFTFMFADPHGKRVFSMLETLYDFIOVHKDDL-QDWLFVLLTOLKK 774
 QY 930 LGTDILNHSKIKWTLQVNHVEFPTQOLKELFRISSTQPTTKTIAIRFLTDLA 969
 Db 775 WQADLGSVQAVYQXALVDTRSPFNDLOFNIMERTVQOTPSLKVAALIKYETLA 834
 QY 990 NTRYKSSDPSPD--OSACERTVLKLAQADQKSELSQARSCLVALYNLNTPTMTLL 1047
 Db 835 ---KQMD-PGDFINSSFRRLAVSRVITWTTERKSPDVAKAQSVALISLPELNTPEFTWL 869
 QY 1048 LADLPKVYDASRCSHSHMRQOSCSGANSPPSSSPK-----PLGSPSYG 1100
 Db 890 LGLLPTFGDGAATKLHNHLR---NTGNGTOSMSGSLTRPTPRPAMWSLTPTYT 945
 QY 1101 PRASLQSHHQLSISTSRSSQSV-----QELPSESLDIQNIKTSSE---INH 1151
 Db 946 SQNTLSPSAFYDVTENWSEDIYSSLRGVTAIQLNPFSSQDEWNEPLRDKSKDDGDSDM 1005
 QY 1152 CFGGQYGTALADNGFNGLQYHDQGDSCASLSSNKTOSGANTTOSN---TPESATM 1207
 Db 1006 CGG-----PMSDPRAGGADTSSQATLADNKLKSLHSMPTHSSPRBRDVPYVYSDSISPF 1061
 QY 1208 RLDNLERERTONAKSPTDADKAVITVSNMANGELLASNLMESEVVRVALTLTKDQPV 1267

Db 1062 NKSALKEAMFPDDADQFPDD-----LSLDHSD---LVAELIKELS-----NHMERV 1104
 Qy 1268 ELLQTSILNIGICIGKNCGLPNKHSRSMMLNLTLEAETHDVYIAGLHYLSKIMRSNK 1327
 Db 1105 EERKIALTELKMLTQESBSFVWDEHFKTLLLETLDGKEPTIRALALKVRLERLHQ 1164
 Qy 1328 MRHMMHLELEILLKTIQCYOH-SKEALRDISMIRIAPSLPLDLSINIVPIATGEF 1386
 Db 1165 AR--FKYAVELTVKMLTEAHDOPHKEVVRSAEASVLATSIPEQCIXVLCPIIOTADY 1222
 Qy 1387 PTNLCAIKILLVEVHEHSEITDAHLIVFPLNASADDTQSMKRAVFCVLKLYFTLG 1446
 Db 1223 PINILAAIKQTKVIERVSKETLNLPLPEIMPGLIOGYNSESVKACVCLVAVHAVIG 1282
 Qy 1447 EEKVKPKSVLPNSKVRLLNVTIEKORNCISGGSGSTKNSAAS 1490
 Db 1383 DE-LKPHLSQLTGSKMKLNLTYI-KRAOTSGGADPTTDVSGQS 1324

RESULT 6
 AAM39287
 ID AAM39287 standard; protein; 1295 AA.
 AC AAM39287;
 DT 22-OCT-2001 (first entry)
 DE Human polypeptide SEQ ID NO 2432.
 XX Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN 26-JUL-2001.
 PD 26-DEC-2000; 2000WO-US034263.
 PF 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00662191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HYGE-) HYSEQ INC.
 PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QH;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX MPI: 2001-442253/47.
 DR N-PSDB; AAI58443.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Example 4; SEQ ID NO 2432; 10078bp; English.
 PS The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nocotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activites such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

XX Sequence 1295 AA;
 SQ
 Query Match 17.7%; Score 1344.5; DB 4; Length 1295;
 Best Local Similarity 29.8%; Pred. No. 1.2e-91;
 Matches 404; Conservative 237; Mismatches 523; Indels 191; Gaps 40;
 Qy 237 SALK-----NTNGNGVGLDEADNIGLRPRPTMIRKPLHSVSSLSLRPPNNV-DVTGDAG 291
 Db 31 SAFKVPAPKTSGN-----PANSARKP-----GSAGGFVAGASKEGAG 70
 Qy 292 AVTWESPESSPEVVPOLNIFPAKMDIYKQVLVIISDKNADMEKRVDAKKIRALL-- 349
 Db 71 AVDEDDFIKAPTDVPSIQYSSRELLETINKIREILSDDKHDMQORANALKIRSLVAG 130
 Qy 350 -LSYHTQPFVAVQKELSLSFVDILKEELASOVIREFACTITAVMSKTLRNKLDAPCWSI 408
 Db 131 AAOYDCFPQHLRLDGLAKLS-----ADLNSOVREACTIVVHLSTVGNKPFHGAGAI 185
 Qy 409 LEHLINLIONSAKYIASASTIALKYIIKYTAAPKLYITDTLNOSSKDIRSTLCMLV 468
 Db 186 VPTLFNLVPSAKWATISGCAIRIIRHTVPRILPITSNCT-SKVVPVRRFEFFLD 244
 Qy 469 LTFEEMQTKALERNATVLRDTLKKSIGDADCDARRHSFYAYAFRRHPELADQYGLTD 528
 Db 245 LLLQEMQWHSLEHNAVLVETIKKGIHDADAEARVEAKTYGMLNHPGEKETLYNSLE 304
 Qy 529 IAAQALEREKREGGGGGTGTGTGAPETRRTVSRIKPTGLQKPTSMRISAVDFPA 588
 Db 305 PSYQKSLQTYLKSAG-----SVASLPQSDSSS---SQESLNPFPSSKMTANPSTVA 355
 Qy 589 ----AQRAKVAQVLYTSRQKPLGPNNSNQASMTGAAA-----SGSLPRPRLNNSGCT- 639
 Db 356 GRVSHGSSKASSLPSSLORSASDIDVNAAGAKAHNAAGSVRSGRIGALNAGSYASL 415
 Qy 640 -----PATTPGSVTPRRGRAG---VSQSPGSSSTSP-----STKLADQYGIEN 682
 Db 416 GRVRAKLAPLAGMNAADASGRSRTYKVVQSQSGSRSGPGRVLTTLALSTVSSGVOR 475
 Qy 683 YVRGATGAI PKKASGIPSTASSRSTSPR---SGGGLMKRSMYSGA---GSRRTPEHN 736
 Db 476 VLWNSASA--QKRSXIPRSGGCSREASPSRLSVASRSRIPRVSQSGCSREASRESRST 533
 Qy 737 NVRPSAP-----ARLDAQSEAEHTLGVGDGQ 766
 Db 534 SPVRSFQPLASRHSRSGALYAPVYASGCGYIGSSSRSLSSVSARVINTGSD--V 591
 Qy 767 DVVSGDYMSGGMGRK-----LMGRESDDIDSEASVCSERSFPSSYTRKXKSNY 819
 Db 592 EBAVADALLGDIRKKYKPARRYESYGMHSDDDANSDASACSESRYSVS-----RNG 644
 Qy 820 SLGSHTRLDNSTORAPFDITIIQFCASSTMSEKDLGLISLTVLAQCKELTQOOLKC 879
 Db 645 SI-----FTYKQTEVAEVLNRCASSNMSEKBLGLQNLKXORTLSREIKR 695
 Qy 880 VLDMEKXEMDQHTK-VYSLFDTVETELIVANETSRNGSSCLTRLFNKLTGDLNLSM 938
 Db 696 LCEIFTRMPADPHGRKRVSMFLFTLVDFIOVHKDDL-QQMLFVLLQOLKKMGADLLGSV 754
 Qy 939 HSKWTKQVHVEYFTQOLKELFRISDSGTPTTKTRIALRFLTDLATYCKSSDF 998
 Db 755 QAKVQALDVTESFPNDQFNILRFTVDQOTBELKVAIKVAILKYEITLA-----KQMD- 809


```

QY 999 PSD--QSQACERTVYLKLAQLADQKSMELRSQASCVALYNLTPOMTLLADLPVYQ 1056
D 810 PGDEINSEETRLAASRVITMTTEPKSSPVRAQSVLISLELNTPEFTMLGALPTFQ 869
QY 1057 DSAASCISHHRRROSQSCNSGANSPPSSPLSSSPK-----PLOSPPVCPFASLOSH 1109
D 870 DGAATKLLHNHLR-----NTGNGTSSMGSPLTPTRPSANWSSPLTSPNTSQNTLSPSA 925
QY 1110 HQLSISSTSPRSSVSE-----QELFSSSELDIOHNIQKTSSE--IRHCFGGQYQTA 1160
D 926 FDVYTEMNSEDIDYSLRGVTEALQNFSGEDMNEIKKDKDGDGDCG-----PG 981
QY 1161 LAPNGFNHLYQHDOGGQSCASLSSNSKTQSSANTQSN---TPESATWRLDNLERER 1216
D 982 MSDPRAGDADTDSQTDALDNKASLLHSPWTHSSPRSDYNNYNSDSISPFNKSALEAM 1041
QY 1217 TTQNAKSPPTDAKITYSINNAENGELILASNMESEVVRVALTLTKDQPELLQTSLTN 1276
D 1042 FDDADQFPDD-----LSLDHSD---LVALLKELLS-----NHNERYEERKIALYE 1084
QY 1277 LGCICIKGNCGLPNKHFPSIRMLNLTLEAHTDVIALGLHVLSSKINRSMKRNHMHFL 1336
D 1085 LMKLTQESBSVMDHETKTLILLLETLDGKEPTIRALAKVLEIRHQPAR--FNQYA 1142
QY 1337 ELILKTIQCYQH-SKXALRDIDSMIRIAPSLPLDSINIVNPVATGEFPTNLCAIKI 1395
D 1143 ELTVKTELEAHKDPHEKVSASAEASVLAITSISPECQIKVLCPIIQTADYPIVLAIKM 1202
QY 1396 LLEVTENHGEITDAHLDIYFNPRLABADDTQSVNRKAAFCVLYKLVGEENVKPKLS 1455
D 1203 QTKVIEKVSKEITLLLPETIMPGLIOGYDNSESSVRACVCLVAHVAIGDE-LKPHLS 1261
QY 1456 VLNPSKVRLINVTIEKORNCISGGSSSTKXSSAAS 1490
D 1262 QLTGSKKKLNTLYI-KRAQTGSGGADPTTVSGGS 1295

```

RESULT 7
 AAM39286
 ID AAM39286 standard; protein; 1309 AA.

AC AAM39286;

DE 22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 2431.

Human; nocotropic; immunosuppressant; cyostatic; gene therapy; cancer;
 peripheral nervous system; neuropathy; central nervous system; CNS;
 Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 leukemia.

OS Homo sapiens.

PN MO20015312-A1.

XX 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US034263.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-0048725.

PR 25-APR-2000; 2000US-00552317.

PR 20-JUN-2000; 2000US-00598042.

PR 19-JUL-2000; 2000US-00620312.

PR 03-AUG-2000; 2000US-00653450.

PR 14-SEP-2000; 2000US-00662191.

PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.

PA (HYSE-) HYSQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Qa;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI, 2001-442253/47.
 DR N-PSDB; AA158442.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PS Example 4; SEQ ID NO 2431; 10078bp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nocotropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification

Sequence 1309 AA;

Query Match 17.7%; Score 1344.5; DB 4; Length 1309;
 Best Local Similarity 29.5%; Pred. No. 1.2e-91;
 Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;

```

QY 237 SALK---NTNGNGVGLDADNIGLRERPTKRLPHSAVSSSLRPKRNVDYGDAG 221
D 31 SAKVPAKPTSGN-----PANSARKP-----GSAGPKVAGASKEGAG 70
QY 292 AATMESFESSFEVVPOLNIFHAKMDMDIYKOVLVIRDKNADBEKEDVALLKIRALLI-- 349
D 71 AVDEDDFIRAFIDVBEIQIYSSRELETANKIRELSDKHMDQAMLKIRSLIVAG 130
QY 350 -LSYHTQPOFVAVOLKELSLSPVDILKEELRSQVIREACTITAAVSKTLANKIDAFCSI 408
D 131 AAQYDCFPQHLRLDQALKLS-----AKDLRSQVVEACITVAHLSTVLGNKPDHGAEL 185
QY 409 LEHLTLVIONSAKVIASASTIALKTIYIKTHAPKLIKITYDTLNOKSDIRTELELW 468
D 186 VPTLFWLVNSAKVATSCAARIFRIRTHVPRLLPLTISNCT-SKSVVRRRSFEFLD 244
QY 469 LLEFEHQTALEARNATVLDLTKSIGDCCARRSRRAVYAFRHPFELADQIYGLTD 528
D 245 LLIQEMQTHSLERHAAVLETTIKGIHDDADARVAKRTYVGLRHPFGEAETLYNSLE 304
QY 529 IAAQRALEREREGGG-----GGTGTGTAPETRTVSR 563
D 305 PSYQKSLQTYLKSSGVSALPQSDRSSSSSQSLNRPFSSKWNSTANPSTVAGRVASGSK 364
QY 564 IGRTPGTLOKPTPSKRSISAVPTAAQAQAKVPAQVTLYRQKKPLGPNNNSQAS----- 617
D 365 ASLTPGSLQ-----RSRSDIDVNAAGAKAHAAAGQSVRSGR-LGAGALNAGSYASLED 417
QY 618 ----MTGAAGAGLPPRLNNSNGGTPATTPSVTPRPGRAG---VSGQSPGRSTSP- 669
D 418 TSDKLDGTASBDGRRAKLSA-----PLAGKNAADSGRGRRTKRVVSSQSPGRSGSGC 472
QY 670 ----STKLDQYGGIGNYRGATGAI PKKASGIPSTASRSTSPTR---SGGGLKRSN 722
D 473 RVLTTLTALSTVSSGVORVLNBSA--QKRSKIPRSGGSRBASPSRLSLVASSRIPRPS 530
QY 723 YSTGA---GSRRTPERNNVPSAP-----ARLLA 749

```

```

Db      531  VSGGCSREASRRESRDTSFVRSPFOPLASRHHRSSTGALYAPRYGASGPGYIGSQSSRLS 590
Qy      750  QGRREHLTGVDGDDQDPDVVSQDVMRSGGMRGRK-----LMGRDESDDIDSEASSVC 802
Db      591  SSVASMRVYNTGSD--VEEAVADALLGDIRTKKPPARRRYSYGMHSDDDNSDASSAC 648
Qy      803  SERSPDSSYTRGNKSNYSLSGSHTRLDWSTQAPAPDDEITIIIOFCASTHMERKQGLISL 862
Db      649  SERSSSS-----RNGSI-----PTVMRQEDVAEVLNRCASSWSRKEGLGL 692
Qy      863  TOYLADKELTOQOLKCVLMFRKMPMDTHYK-VISLFDITYTELLVHANTSRNGSSS 921
Db      693  QNLKNGORTSLRVELRICEIFTRMPADPHGKRVSMELETIVDFIOVHKDL-ODWLFV 751
Qy      922  CLTRFNLKGTDLNLSMHSKIKTLOVNEVEFPTQOLLEPRISDSQTPPTKTRIAI 981
Db      752  LITOLLKKGADLLGSVQAKVQKALDVTRESFPNDIQENILKRFVVDQOTPSLAKVYAI 811
Qy      982  LRFITDLANTYCKSDPSPD--QSQACERTVLKLAQLAADOKSMELRSQARCLVALYVL 1039
Db      812  LKRIETLA-----KQMD-PGDFINSSETRLAVGRVITWTEPKSSDVKKAQSVLISLFL 866
Qy      1040  NTPQMTLLIADI.PKYVQDSARSCISHMRQSQSCNSGANSFSSSPSSSK----- 1092
Db      867  NTFPEFTMLGALPKTFQDQATKLIHNLH-----NTNGTQSSWGSPLETRPTPRSPAMSS 922
Qy      1093  PLOQSSVGFALQSHHHLQSSISSTSPRSROSSVE-----OELPFSSELDIOHNIQTS 1146
Db      923  PLTSPNTISQNTLSPSAFDYDTENMNSEDIYSLRGVTAIQWFSRQEDMNEPKRDS 982
Qy      1147  EE---IRHCGGQGYOTALPNGFNGLQYHDOQODSCASLSNSKTQSANTTQSN--- 1200
Db      983  KQDDGDSMGCG---PGMDPRAGGATQSSQFALDNKXSLHSMPTSSPSRDNPNYN 1038
Qy      1201  -TPESATMBLMDLERERTQNAKSPTDKAVITVSINMAENGELIASULMESEVVRVAL 1259
Db      1039  YDSISIPFKSALKKAMFDDADQFPD-----LSLDHSD---LVAEELKELKS----- 1083
Qy      1260  TLTQKQPVELLQSLNLGICIGNCELPNKHFRIMMLNLILAEHTDVYIAGLYL 1319
Db      1084  --NNHNRVERKALVLEMLKTQESFSVWDEHFXTLLLTLDKPEPTIALALKYL 1141
Qy      1320  SKIMRSNKRHNMHGLELILKIIQCYOH-SKEALRIDDSMIPRIAPSLPDISINIVN 1378
Db      1142  RETLRHQPAR--PKNYAEVLVKTLEAHDPKHEVRSAEBAASVATISPEQCICKVC 1199
Qy      1379  PVATGEFPTNLCAIKILEVTEHSGSEITDAHLDIVFPNLARSADDTQSMVKAAPCT 1438
Db      1200  PILOTADYRINLAIMQTFVIERVSKETLNLPLPEIMPGLIQGYDNSESSVAKACVFCI 1259
Qy      1439  VKLYFLVGEERKVPKLSVLPKSVRLNVTIEKORNCISGGSGSTKSSAAS 1490
Db      1260  VAVHAIVGDE-LKPHLSQLTGSKKCLNLYI-KRAQTGSGADPTTIDVSSQS 1309

```

RESULT 8
AAM41072
ID AAM41072 standard; protein; 1317 AA.

AC AAM41072;

DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6003.

XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.

```

OS Homo sapiens.
XX WO200153312-A1.
XX 26-JUL-2001.
XX 26-DEC-2000; 2000WO-US034263.
XX 23-DEC-1999; 99US-00471275.
XX 21-JAN-2000; 2000US-0048725.
XX 25-APR-2000; 2000US-00552317.
XX 20-JUN-2000; 2000US-00598042.
XX 19-JUL-2000; 2000US-00620312.
XX 03-AUG-2000; 2000US-00653450.
XX 14-SEP-2000; 2000US-00662191.
XX 19-OCT-2000; 2000US-00693036.
XX 29-NOV-2000; 2000US-00727344.
XX (HUSE-) HYSEQ INC.
XX Tang YF, Liu C, Aundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX MPI: 2001-442253/47.
XX N-ESDB; AAI60228.
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX Example 2; SEQ ID NO 6003; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AAM3642-AAM42213) with noctropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localized neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemia and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX SQ Sequence 1317 AA;
XX
XX Query Match 17.7%; Score 1344.5; DB 4; Length 1317;
XX Best local similarity 29.5%; Pred. No. 1.3e-91;
XX Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;
Qy 237 SALK-----NTNGNGUGLDADNIGLRERPTMKPIHSAVSSSLPKPNVN-DVTGDAG 291
Db 339 SAKVPAPKTSGN-----PANARKP-----GSAGPXYGAGSKSGAG 78
Qy 292 AVTMESEFSEFVVPQNLIFHAKMDIYQVLIISDNADWEKRVDAIKIRALLI-- 349
Db 79 AYDEDDFIKAFIDVDSIQYSRELEETLNKIRELISDDKHMDQANALKIRSLIVAG 138
Qy 350 -LSYHTQPOFAVQKLSLSFVDILKELRQVIRACITTAAYMSKTLRNLDAFCWSI 408
Db 139 AAQYDCFFQHLRLILGALKLS-----AKLRSQVAREACTIVAAHLSVLGNKFDHGAELI 193
Qy 409 LEHLINLQNSAKVIVASASTIAKLYIKYTHAPKLIKITYDTLQSKSDIRSLTCELV 468
Db 194 VPLFLVFNASKAKWATSGCAIRPIIRHTHPRLIPLTNSCT-SKSVVRRRSPEFLD 252
Qy 469 LTFEEMQTLARNAIVLBDTLKSGDAGCDARRHSRYAYAFRRHPPELADQIYGTLD 528
Db 253 LILQEWQTHSLERHAIVLVEITIKGIHDADEARVARKTYMGLRNHFPGEAETLVNSLE 312

```

```

QY 529 IAAQALREREREGGG-----GCTGTGTATPRTTRYSR 563
DB 313 PSYQSLQTYLKSSGSVASLPQSDRSSSSQESLNRPFSSKMTANPSTVAGRVASGSK 372
QY 564 IGRTPGTLOKTPPSKRSISAVDTAAQAKVRAQTYLTSRQKPLGPNNSQAS-----617
DB 373 ASSLPGSLQ-----RSRSDIDVNAAGAKAHHAAGQSVRSGR-LGAGALAGVASYLSD 425
QY 618 ----MTGAAGSLRPRPLNSNGTTPATPGSVTPRRGRAG---VSQSPGSRSTSP- 669
DB 426 TSDLDGASDEGRRAKLSA-----PLAGMNAKADSRGSRMTMSQSPGSRSGSPG 480
QY 670 ----STKLRDYGGIGNYRGATGAI PKKASGIPRSTASRETSPTR---SGGLMKRSM 722
DB 481 RVLTTTALSTVSSGVQVLVNSASA--QKRSKIPRSGCSSEAPSRSLSTVARSRI RPPS 538
QY 723 YSTGA---GSRTPERRNPPRPSAP-----ARLLA 749
DB 539 VSQCSREASRSDTSRPSRFPOLASRHSRSTGALYAEVYAGSGPGYGISQSRSL 598
QY 750 QSRBAHTLGVGDGQDPYVSGDYMRSGGMGRK-----LMGRDESDIDSEASVC 802
DB 599 SSVAMRYLNTGSD--VEEAVADALLDITRKPPARRRYESYGMHSDDANSASAC 656
QY 803 SERFSDSYTRGNKSNYSLSGHTLMDSTORAFDIDETIICFCASTHMERDGLISL 862
DB 657 SERYS--RNGSI-----PTMRQTEDEAVELNRCASNMEREGGLGL 700
QY 863 TQYIADGKELTQOOLKCVLDMFRKKFMDTHK--YSLFLDTVTLLVHANETRNSSS 921
DB 701 QNLKNQRTSLRVELKRLCEIFTRMFADPHGRKVFMSFLETLVDFIQVHKDDL-QDMLFV 759
QY 922 CLTEFLNLTGTDLNSMSKIMKTLQVHEXEPQLOLKEFLRISSTOTPTKTAT 981
DB 760 LLTQLLKKMGADLLGSVQAKVQKALDVRESFMDLQNLIMRTVDOTVDTSLKVKAI 819
QY 982 LRFLTDLANTYCKSSDFPSD--QSOACERTVLKLAQLADQSMELNSQARSCVALYNL 1039
DB 820 LKYEITLA---KQMD--PGDFINSETRLANVRITWTEKSDVRAQAQSVLISL FEL 874
QY 1040 NTPQMTLLADLPKYQDSARSCSHMRROSQCSNCSNPPSSPLSSSPK-----1092
DB 875 NTPFTMLLGLPFTFDGATKLLHNLH---NTGNQTSMSQSPLTRPPRPSANWSS 930
QY 1093 PLOGSVCFASLQSHHQSLSSSTPRSSQSV-----QELFSEELPQNIQIKTS 1146
DB 931 PLTSPNTSONTLPSAFDYDTENNSEDIYSLRGVTEALQNFSPSQEDMNEPLKRS 990
QY 1147 EE---IRHCFGQYOTALPANGFNGLQYHDQGOODSCASLSSNSKTQSSANTTQSN---1200
DB 991 KKDGDGSMCGG---PGMSDPRAGDADATDSQTLADNKASLLHSMPTTSSPRSDYPN 1046
QY 1201 -TPSSATKRLDNERERTQAKSPTDQAKYITVSINAEENGELILSNLMESEVRYAL 1259
DB 1047 YSDSISPNKSALEAMFDDADQFPD-----LSLHSD---LVALLKELS-----1091
QY 1260 TLTQDQPELLQTSLTNLGICIKGNCCLPKHRSIRMLNLEHNDVYAGLHYL 1319
DB 1092 --NHNERYEERKIALYELMKLTQEESSFSVWDEHFTLLLETLTGDEPTRLALKVL 1149
QY 1320 SKIRSNMRRNMHMFLELLIKLIIQCYOH--SKELAPDIDSMIPRIAPSLPLDSINIVN 1378
DB 1150 REILRHQAR--FNKVAYELTMKTLKLEAHKDPHKEVVRGAEBAAVLAATISIEPCIKVLC 1267
QY 1379 PVIAITGEPTNLCAIKILEVTEHHGSEITDAHLDYFNPILARSADTQSNVRGAANFCI 1438
DB 1208 PIITADYPIIMAIKQOTKIEEVSXETLNLILPEIMPGIIOGDNSESSVRAACVFCI 1267
QY 1439 VKLFFVGEKVKRSLVNSKVRLLNVEYKORNCISGGSSSTKUSSAAS 1490
DB 1268 VAVHAIVGDE-LKPHLSQLTGSKKLLMLYI--KQAQTSGGADPTTVDVGOS 1317

```

```

RESULT 9
AAM41073
ID AAM41073 standard; protein; 1317 AA.
XX
AC AAM41073;
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 6004.
XX
KW Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
peripheral nervous system; neuropathy; central nervous system; CNS;
Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
leukemia.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUL-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Qa;
PI Zhou P, Goodrich R, Dmanac R;
XX
DR WPI; 2001-442253/47.
XX
DR N-PSDB; AAI60229.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
PS Example 2; SEQ ID NO 6004; 10078bp; English.
XX
The invention relates to human nucleic acids (AA157798-AA161369) and the
encoded polypeptides (AAM38642-AAM42213) with nocotropic,
XX
CC immunosuppressant and cyostatic activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
XX
CC system, such as peripheral nervous injuries, peripheral neuropathy and
central nervous system diseases, such as
XX
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX
CC utilisation of the activities such as: Immune system suppression,
chemotactic/chemokinetic activity, haemostatic
XX
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
assays for receptor activity, arthritis and inflammation, leukaemias and
XX
CC C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX
SQ Sequence 1317 AA;
XX
Query Match 17.7%; Score 1344.5; DB 4; Length 1317;
Best Local Similarity 29.5%; Pred. No. 1,3e-91;
Matches 405; Conservative 233; Mismatches 523; Indels 211; Gaps 40;
QY 237 SALK---NTNGNGVGLDEADNIGLREPTRMKIRPLHSAVSSSLRPKRVN-DVTGDAG 291

```

```

Db      39 SAFKVPAPKTSGN-----PANSARKP-----GSAGGPKVAGAGSKGGG 78
Qy      292 AATMESFESEFVVPOLNIFHAKMDIYKQVILISDNADBEKXVDLKKIRALLI-- 349
Db      79 ADEDEDFIAFDVPSIQIYSSRELEETLNKIREILSDGHDQDAALKKIRSLVVG 138
Qy      350 -LSYHQPOQFVAVOLKELSLSPVDILKEELRSQVIREACITIAVMSKTLRNKLDAFCWSI 408
Db      139 AAQYDFFQHLRLDGLAKLS-----AKOLRSQVREACITIAVMSKTLVGNKFDHGAEL 193
Qy      409 LEHLINLONSAXVIASTIALKVIKITHAPKLIKITYDTLNOSKSDISTCELV 468
Db      194 VPLFLVFNPAKVMATSGCAIRFIIRHVPRLPLITSNCT-SKSPVVRRSFEFLD 252
Qy      469 LFEEMQTALENANVLVDLTKKSGDGDCCDARRSRAYAFRRHPELADQIYCTD 528
Db      253 LILQEMQTHSLERHAAVLVETIKGIIHDADAEARVEARKTYMGLRNHPGEMLTLYNSLE 312
Qy      529 IAAQRLEREREGGG-----CGTGTGTGTAPETRRTVSR 563
Db      313 PSYQKSLQTYLKSSGVASLPQSDRSSSSQESLNPFSKMTANPSTVAGRVASGSX 372
Qy      564 IGRTPGLQKPTPSMKSISAVDTAAQRAVRAQYTLYSRQKPLDGPNNNSQAS----- 617
Db      373 ASGLPGSLQ-----RSRDIDVMAAAGAKAHHAAGOSVSGR-LGAGALNAGSVASLED 425
Qy      618 ----MTGAASGSLPPRLNSNGGTPATTPGVSITRRPRRAG---VSOSOPGSRSTSP- 669
Db      426 TSDKLDGTASDEDRVAKLSA-----PLAGMGNAKADSRSRSTKVSQSPQSRKSGP 480
Qy      670 ----STRKRDQYGGIGNYRGATGALPKKASGIPRSTASSRERSPTP---SGGGLMKRM 722
Db      481 RVLTTALSTVSSGVQRVLVMSASA--QKRSKIPROGCGCRASPRSLVSRASRIIPRS 538
Qy      723 YSTGA---GSRRTPEPNNPVRPSAP-----ARLLA 749
Db      539 VSGGCGRERASRESRDTSPVRSFQPLASRHRSSTGALVAFVYAGSAGPGYGISOSRSL 598
Qy      750 QSRERAEHTLGVDDGCPDYVSGDYMSGGMKRMGR-----LMGRHSDIDISEASVVC 802
Db      599 SSVSARKVAVMTGSD--VEEAVADALLGDIRTKKKPPARRRYESYGMHSDDANSASAC 656
Qy      803 SESSPSTYRGMKSNVSLSGSHRLDWSIORAPFDIETIIQFCSTHWSERKGLISL 862
Db      657 SESSYS-----RNGSI-----PTMROTEVAVALNRCASSMSEKKEGLGL 700
Qy      863 TOYIADKEITLQOOLKCVLDMFRMFMNDHTK-VYSLFDVTVELIVHANETSRNGSS 921
Db      701 QNLKNGRTLSRVELKRLCEIFRMFADPHGKRVFSMFLTLVDFIQVHKDDL-QDWLFV 759
Qy      922 CLTRLENKGLDILLNSHASKIKWTLQVYVHEFPYQLOKELFRISISTOTPTTKRIAI 981
Db      760 LITQLAKKMGADILGVOAKVQKALDVTBSPFDLONILMRTVVOGTOTPSLKVAVAI 819
Qy      982 LRFLITLANTYCKSSDFSPD--OSQACERTVLKLAOLAQOKSELSQARSCVAVLYNL 1039
Db      820 LKYLETIA---KQMD-PGDFINSETRLAVSRVITMTTEKSSDVYKAAQSVLISLPEL 874
Qy      1040 NTPOMTLIADLPKYQDSARSCISHMRROSCNSGANSPPSSPLSSSPK----- 1092
Db      875 NTPFTMLLALPPTPDGAKTKLHNLH-----NTGQATQSMGSPLTRPFRPAMWS 930
Qy      1093 PLOGSPVGPASLGSHHQLSISTSPRSQSVY-----QELFSELDIONIKTS 1146
Db      931 PLTPPTVTSQNTLSPSAFYDYTEMNSEDYISLRGVTALONFSPSQEDMNEPLKQDS 990
Qy      1147 BE---IHRCGGQYQTLAPGFGNHLQYHDQGDSCASLSNSKTSQSNANTQSN--- 1200
Db      991 KKDDGDSMGCG---PGMSDPRAGDADTDSQTLADNKASLLHSMPTHSSPRSDVYPYN 1046
Qy      1201 -TPESATWRDLNLERERTONAKSPTDPAKVIIVSINNAENGELILASNMESSEVAVAL 1259

```

```

Db      1047 YSDSISPPNKSALEKAMPDDADQPPD-----LSLDHSD---LVAEELKELS----- 1091
Qy      1260 TLTQDQVELLQTSITNIGTICIGKNCGLPNKRPISIRMLNLTLEAHTDVINGIHLV 1319
Db      1092 --NHERVEERKIALYELMKLQESFSVWDEHFTILLTLETTGDDKEPTRALAKVL 1149
Qy      1320 SKIMSNMGRRNWHMFLLEILTKITQCYOH--SEKALRIDSMTPIRAPSPLDLSINIVN 1378
Db      1150 REILHQPAR--FKNYAEITWTKTEAKNDPHEKVEYRSEAEASVLAISISPEQCIKVL 1207
Qy      1379 PVATGEPTMLCAIKILLEYTEHNSSEITPAHLDIVFPNLRASADDTQSMYKAAVPT 1438
Db      1208 PIIQIADPINALIKMQTKYERYSKETLNLLEPEIMGGLIQGYDNSSSSVKACVFL 1267
Qy      1439 VKLIVLGEKVKPKLSTVLPNSKATLVNLYIEKQNCISGSGSSSTKNSAAS 1490
Db      1268 VAVHAVIGDE-LKPHLSQLTGSKMGLNLNLYI-KRAQTSGGADPTTVSGQS 1317

RESULT 10
ADB64143
ID      ADB64143 standard; protein; 916 AA.
XX
AC      ADB64143;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Human protein encoded by clone CTONG20032930.
XX
KW      Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
KW      cell regeneration; membrane protein; signal transduction-related protein;
KW      transcription-related protein; osteoporosis; neurological disease;
KW      cancer; tumour.
XX
OS      Homo sapiens.
XX
PN      EP1308459-A2.
XX
PD      07-MAY-2003.
XX
PF      28-MAR-2002; 2002EP-00007401.
XX
PR      05-NOV-2001; 2001JP-00379298.
XX
PR      25-JAN-2002; 2002US-00350978.
XX
PA      (HELI-) HELIX RES INST.
XX
PA      (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI      Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI      Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI      Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y;
XX
DR      WPI: 2003-450961/43.
XX
DR      N-PSDB; ADB62173.
XX
PT      New polynucleotides and polypeptides, useful for developing a diagnostic
PT      marker or medicines for regulation of their expression and activity, or
PT      as targets of gene therapy.
XX
PS      Claim 1; Page; 222pp; English.
XX
CC      The invention discloses a polynucleotide comprising a sequence selected
CC      from 1970 fully defined nucleotide sequences which encode novel
CC      polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC      or its partial peptide, an antibody binding to the polypeptide or peptide
CC      of the polynucleotide, immunologically assaying the polypeptide or peptide
CC      peptide of the polynucleotide by contacting the polypeptide or peptide
CC      with the antibody of the encoded protein, and observing the binding
CC      between the two, a transformant carrying the polynucleotide in an
CC      expressible manner and an antisense polynucleotide. The oligonucleotide
CC      is useful as a primer for synthesising the polynucleotide, or as a probe
CC      for detecting the polynucleotide. The polynucleotides and encoded
CC      proteins are useful as pharmaceutical agents and many disease-related

```


transcription-related proteins, disease-related proteins and genes encoding them can be used as indicators for diseases (e.g. osteoporosis, neurological diseases, cancer, tumors). The cDNA may be used to regulate the activity or expression of the encoded protein to treat diseases. The sequence presented is a protein of the invention. Note: Some of the sequence data for this patent is not represented in the printed specification, but is based on sequence information supplied by the European Patent Office.

Sequence 723 AA:

Query Match 12.1%; Score 920; DB 7; Length 723;

Best Local Similarity 32.0%; Pred. No. 6e-60; Matches 243; Conservative 144; Mismatches 274; Indels 98; Gaps 19;

5 KPESDLOFIOMPRADNRKVQVLAEDLVTL-SDDTNSIYCTDMGFL---IDGIMPLTGS 60-
2 EPRSMETFCAGVOQKDGRLQVQGLVYGAAGALSDEBEDGLGKIVDALTGWVS 61
61 SHFKIAKSLAEFSELIKRLGSDENATATVLPVHIDRLGDSRDVREKAQLLRDIMEH 120
62 SNRVSLMGLEILSAFDRILSTRKSYAAVVALIDRMGDADKXVDEAQTILIKLMD- 120
121 RVLPQALIDKATSCFHKAKVREEFLOTIVNAHEGTQQLSVAVYIPVCALLGDP 180
121 OVAPPMYIWEQLA-SGFKHKVFRSEGVCLLETLLNIFGAQPLVISEKLPHLCILFSDS 179
181 TVAVREAIQTIVYKGVGDRRLRPMDDVPASLAMEQFEDVQKQEGILLPSALK 240
180 NSQVRDAIILAIYEIVHVEKVMDLKYR-GIPPALEWIFAFEDVQSSGGMILSVCK 238
241 NTGNGVGLDEADNIGRERPT-----RMIRPLHSAYSSSLRKPENVNDVT-- 287
239 DKS-----PDDESV-D-GNRPSAASAFAKVPAPRTSGNPNASAKPPSAGSPKVGASKE 292
288 GDAGAVTMSFESSFEVVPOLNIFHAMDMDIYQVLIISDKNADNEKRVDAKLRAL 347
293 GGAGAVDEDDFIKFTDVPSIQIYSSRELETLNKIREILSDDGHDWDQANALKKIRSL 352
348 LI--LSYHPOPOVAVQKELSLSPFDILKEELRSQIVRACCTIVMSTLBNKLDALF 404
353 LVAGAAQYDCFFOHLRLDGLALKLS---AKDLRSQVREACTIVHLSTVLGNKFDHG 407
405 CWSLIEHLINLIQNSAKVIASASTIALKYIITYHAPKLEKIYDTLNOSKXDIRSTLC 464
408 ABAIVPTLPIVLPNSAKVATSGCAALRFTIRHIVRLLPLITSNCT-SKSVPRRSF 466
465 ELWVLFEEWQTKALERNATVLRDTLKKSIGDADCDARRHSRYVYMAFRHPELADQIY 524
467 EPLDLLLOEQMTHLEHAAVLVETIKGIHDADAARVEARKTYMGLRNHPGAEETLY 526
525 GTLDIAAQRALERREBEGCG-----GGTGTCTGAPETRR 559
527 NSLEPSYOKSLQTYLNSSGVYASLPQSDRSSSSSQESLNPFSKSKWSTANPSTVAGRYSA 586
560 TVNSRIGTPGLTKPTSMRSISAVDPAARAKVRAQYTLYSRQKRLPGNNNSQAS-- 617
587 GSSVASSLPESLQ-----KRSRSDIVAAAGAKAHAAQGVSSGR-LGAGALNAGSYA 639
618 -----MTGAAASGSLPRRLNSNSGCTPATTPGVSVPRRPAG--VSQSQPSGRS 666
640 SLEDTSDLDGTAEDGKVRALISA-----PLAGGNKAKDSRRSRSTXWVSQSPGSR 694
667 TSPTKLRDYGIGGNYRKGATGALPKKASGIPSTASS 705
695 GSPGRVLT-----TTTALSTVSSGVQRLVNS 720

RESULT 12
ADU20651
ID ADU20651 standard; protein, 1330 AA.
XX
AC ADU20651;

13-JAN-2005 (first entry)
A. thaliana At2g20190 homologue.
plant; tolerance; resistance; environmental stress; drought.
Arabidopsis.
thaliana.
MO2004092349-A2.
28-OCT-2004.
15-APR-2004; 2004WO-US011887.
15-APR-2003; 2003EP-00008079.
01-AUG-2003; 2003EP-00016671.
30-SEP-2003; 2003EP-00022226.
(BADI) BASF PLANT SCI GMBH.
Puzio P, Chardonene A, Chen R, Puente P;
WPI; 2004-766956/75.
N-PSDB; ADU20650.
New transformed plant cell with altered metabolic activity compared to a corresponding non transformed wild type plant cell, useful for producing, screening and breeding plants with increased tolerance to environmental stresses.
Disclosure; Page 598-599; 607pp; English.
This invention describes a method resulting in a novel transformed plant cell which has been altered by an inactivated or down-regulated gene and results in increased tolerance and/or resistance to an environmental stress as compared to a corresponding non-transformed wild type plant cell. The method can be used to generate or screen for seeds or plants with increased tolerance/resistance to environmental stress and for detecting stress in cells or plants. The transformed plant cell comprises a metabolic activity that is altered concerning one or more metabolites selected from 2,3-dimethyl-5-phytylquinol, 2-hydroxy-palmitic acid, 3,4-dihydroxyphenylalanine, 3-hydroxy-palmitic acid, 5-oxoproline, alanine, alpha linolenic acid, alpha-tocopherol, aminoacetic acid, abhydroglucose, arginine, aspartic acid, beta-apo-81 catotenal, serine, beta-carotene, beta-sitosterol, beta-tocopherol, hexadecatrienic acid, margarinic acid, gamma-aminobutyric acid, (delta-7-cis,10-cis)-hexadecadienic acid, delta-15-cis-tetracosenic acid, ferulic acid, campesterol, cerotic acid (c26:0), citrulline, cryptoxanthine, eicosenoic acid (20:1), fructose, fumarate, galactose, gamma-tocopherol, gluconic acid, glucose, glutamic acid, glutamine, glycerate, glyceraldehyde, glycerol, glycerol-3-phosphate, glycine, homoserine, inositol, isoleucine, isomaltose, isopentenyl pyrophosphate, leucine, lignoceric acid (c24:0), linoleic acid (c18:2 (c9, c12)), luteine, lyopene, malate, mannose, methionine, methylgalactofuranoside, methylgalactopyranoside, methylgalactopyranoside, palmitic acid (c16:0), phenylalanine, phosphate, proline, putrescine, pyruvate, raffinose, ribonic acid, shikimate, sinapine acid, stearic acid (c18:0), succinate, sucrose, threonine, triacontenoic acid, tryptophane, tyrosine, ubiquinone, udp-glucose, valine, and zeaxanthine. The plant is selected from maize, wheat, rice, oat, triticale, rice, barley, soybean, peanut, cotton, rapeseed, canola, manihot, pepper, sunflower, flax, borage, safflower, linseed, primrose, rapeseed, turnip rape, teagates, solanaceous plants, potato, tobacco, eggplant, tomato, Vicia species, pea, alfalfa, coffee, cacao, tea, Salix species, oil palm, coconut, perennial grass, forage crops and Arabidopsis thaliana. The environmental stress is selected from salinity, drought, temperature, metal, chemical, pathogenic and oxidative stresses, or their combinations. The inactivation or down-regulation of the gene is achieved by double-stranded RNA interference (dsRNAi), introduction of an antisense nucleic acid, a ribozyme, an antisense nucleic acid combined with a ribozyme, a nucleic acid encoding a co-suppressor, a nucleic acid encoding a dominant negative protein, DNA- or RNA- or protein-binding

CC factors targeting said gene or -RNA or -proteins, RNA degradation
 CC inducing viral nucleic acids and expression systems, systems for inducing
 CC a homologous recombination of the genes, mutations in the genes or their
 CC combinations. This sequence represents a homologue of an Arabidopsis
 CC thaliana drought tolerance protein.
 XX

Sequence 1330 AA;

Query Match 8.0%; Score 608.5; DB 8; Length 1330;

Best Local Similarity 20.1%; Pred. No. 5.6e-36;

Matches 295; Conservative 290; Mismatches 618; Indels 267; Gaps 52;

122 VLPQALDKATSCCKKNAKREBFLOTIVNALHEVGTQOLSV-RVYIPVYCALLGDP 180
 3 VSSPIIVERASYSMMHSMRVRREFPARTVSAIGLFSTELPLQRIVLATILMLNDP 62
 181 TNNVEAIIQTLVEYKIVDRLPDLRMDVPASKAMLEQKDOYVQESLELLPSALK 240
 63 NOAVEAALICEEMMGQSQFRELQK-HHLPSIMVADINARLERLEPQ-----LR 114
 241 NTNGNGVGLDEADNIGLRERPTMIKRPILHSAVSSSLRPKN-----VNDVTGDAG 291
 115 STDG-----RSAAHVYNEVKASSVNPCKSPRAKAPTRNSLFGDA 156
 232 AVTMSESSFEVNPQNLIFHAKDM-DDIYKQVLIISDKNADWEKRVDAIKIRALLIL 350
 157 DITEKPIE-----PIKYSEKELIREPEKIAATLVEPK--DMGMRISAMRVEG-LVA 206
 351 SVHTQPOFAVQKELSLSFVDILKEEL--RSOYIREACTIATMSKTLRKLDAFCMS 407
 207 GGATVYSCFRLKQKQVLPSTQADRRSTVYKQACHLLCLLSKELLGDFPACAE 262
 408 ILEHLINILIONSAKVIASASTIALKYIIKYTHAPKLKLYDTLNOSSKDIRSTCELM 467
 263 FIPVFLKLVITVIAESADNCIKTMLNCKAAVLPRIASAGHDNALIRACCCEYA 322
 468 VLLFEEM-OTKALEFNATVLRDTLKKSIGDADCDARRHSRYVYAFRRHPELADQIYGT 526
 323 LITLHEWPDAPFIQKSVLDYEDLIRCCVADAMESEVATAMCYRMFAKTWPDERSRLPSS 382
 527 LDIAORALEREDEGGGCTGTGTGTAETRTATSRIGRTCTLOKPTPSMRISAVDT 586
 383 FDPVIGRLINE---DGIHRRHASPSYREHSHQPSFSQTSAPSNLPEYGTSAIVANDR 438
 587 AAACRAKRYAOYTLXSRQKPLGPNNSN-----QASMTGAAGSGLPR-----P 630
 439 SSNLSGGSLSSGLLSSQKDVNKGSESLSVLSSQKXSAISMLRGHLISDRNPA 498
 631 RLSNSGGTPTTPGCVTPPRGRAGVSGSQPS---RSTSPSTKLDOYGGIGNYRGA 687
 499 ALRSSSLDLGVDPSPSRDPPFHAIVAPASNSHTSSAAESTHSINKSNRNGGLG----- 552
 688 TGAIRKXASGIPRSTASSRETSPTSSGGLMKRMYNS-TGAGSRTPPRNNVVRSAAR 746
 553 -----LSDIITQIQASKDSGRSSYKGNLSESHPTSSLTAKKGSERNE--RSS--- 599
 747 LLAQREAEHT---LGVDDGQDPYVSGD--YMSGGGMR-----GRKLING 787
 600 -LEESNDAREVAFMAAGHFDROQMDTAYRDLTFRSMNSHVNFORPLLRKXVGRMSAG 658
 788 RDESDIDSEASSVCSERSFSSSYTR-GNKSNY-----SLG-----GSHTRLDWSTORAP 836
 659 R-----RRSFDSQIQIDISNFGVGPASLINALDGLNSSSDWCARVAA 703
 837 FDDITFIIOFCASTHMSERKQGLISLTQVLADGKELTQOQLKCVLDMRKXMPMDHTHY 896
 704 ENFLDTLLQ-----QG-----PKGAQEVITQSEFKYKVLFLRHLDHNNHYA 744
 897 SLFLDTVELILVHNETSRNGSSC-----LTRLFNKLGTDLLNSMHSKIWT 945
 745 QAALSTLADLI-----PSCRKPFESYMERVLPHVSRLL-IDPKEVVRPCGST 791
 946 LQVNEHYFPQOLQKELFRILISDSQTPTTKTRIALILRFLDLANTYCKSSDFPSDQOA 1005

DB 792 LEIVSKTYSDSLPALRLSL-DEQRP--YAKAVIEFAINSFNRYAGNEI-SGNSGI 847
 QY 1006 CERVLTKLAQADQKMELESCARSCIALYN-LNTPQMTLLADLPKVYQDSARSGIH 1064
 DB 848 LKMLAKTLPLTRD-KYTKLKEASITCLISVYNNHYDAGLNYLISLVEEONGLRRLX 906
 QY 1065 -----SHNRQSQSCNSGANSPPSSPLSSSPKPLQSPS-----VGPRF-SLQ 1106
 DB 907 QYTRIEVDLLNYQSKKQKORIKSYDPSDA-IGTSSEGYAGASKXNIFLGRYSGSID 965
 QY 1107 SHHQLSISSTSPRSRQSVQELLFSELDIQNNIK-----TSEIRHCFGGQY 1157
 DB 1158 QTAL---APNGFNGLYHQGGQDSCASLSNSKTSQSANNTQSPESATMELDNLER 1214
 1026 QNSISRTSPNOSSENIILD---DLSPHLEKQNLNTSVDSLGRHENEVSRELDGHY 1082
 QY 1215 ERTTQNAKSPTDKAVITVSINMAENGELILASLMESEVVRVALTLTKQPVELLQTS 1274
 DB 1083 MLTSIKVNTTPESGPSIPQILHMINGSDSPSS-----KSGIQQLIEASV 1129
 QY 1275 TNLGICIKGNCGLPNKGFPSIMMLNILEAETTDVVIAGLHLSKIMRNNK--MRNNW 1332
 DB 1130 AN-----EESWTYTFNQILTVLEVLDEDEFSIKELALSISEMLKSKQKAMEDS- 1180
 QY 1333 MHFELILKLIQCYOH-----SKALRDIDSMIPRIAPSLPLDSINIVNPVIATGEP 1387
 DB 1181 ---VEIYEKLIHNSKQTPRVNSTEABOCLTVVSQYDF---RCLSVIPLLVTEBEK 1233
 QY 1388 TNLCAIKILLEVTEHSGSEITDAHLDIVPNLARSADTQSMVRAVFCIVLYFVLGE 1447
 DB 1234 TLVACINCLTFLVGRLSQBELMDQLSSFLPAVFAFGSQSADYKTVVFCIVDIYIMLGK 1293
 QY 1448 EKVPKSLVNPSTKRLNLY---IEKQRN 1474
 DB 1294 AFL-PYLEGNSTQVRLVLTIANRISQARN 1322

RESULT 13
 ADB63929
 ID ADB63929 standard; protein: 454 AA.
 XX
 AC ADB63929;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human protein encoded by clone BRACE20079370.
 XX
 KW Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 PN EPJ308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 PR 28-MAR-2002; 2002EP-00007401.
 XX
 PR 05-NOV-2001; 2001JP-00379298.
 XX
 PR 25-JAN-2002; 2002US-00350978.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S,
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuno Y,
 XX

DR WPI; 2003-450961/43.
 DR N-PSDB; ADB61959.
 XX
 PT New polynucleotides and polypeptides, useful for developing a diagnostic
 marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 XX
 PS Claim 1; Page; 222pp; English.
 XX
 CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.
 XX
 XX Sequence 454 AA:
 SQ
 Query Match 6.6%; Score 502.5; DB 7; Length 454;
 Best Local Similarity 32.5%; Pred. No. 1e-28;
 Matches 138; Conservative 77; Mismatches 156; Indels 54; Gaps 12;
 QY 261 PTRNKRPLHSAVSSSLRPKNVNDVTGDAGAVTMESSEFSEVVPOLNTHADMDIY 320
 DB 44 PANSARKP-----GSAGPKVGASKEGAGAVDEDFIKAFITVPSIQIYSRELEBTL 98
 QY 321 KQVLITISDKNADMEKRDALKKIRALLI---LSYHTQPOFVAVOLKELISFVDILKEE 377
 DB 99 NKIREIISDDKGDMDQDANALKKIRSLVAGAAGDCCFQHLRLIDGALKLS-----AKD 153
 QY 378 LRSQVIREACTITIAVMSKTLNKLDAFCMSILHNLINLIONSAKIVASASTIAKYLIIKY 437
 DB 154 LRSQVIREACTITIAVMSKTLNKLDAFCMSILHNLINLIONSAKIVASASTIAKYLIIKY 213
 QY 438 THAPRLKIYDTLINSKSKDIRSTLCELMVLLFEEMOTKALERNATVLRITLKKSIGDA 497
 DB 214 THVRLPLILITNSCT--SKSVFVRRRSFEFLDLLLOEMQTHSLERHAALVETIKKGIDA 272
 QY 498 DCDARHRHRYVMAFRHRFPPELLADQIYGTLDIAAQRALEREREGGGGTGTGTAPEY 557
 DB 273 DAARVREARKTYMGLRNHFPGEAETLYNSLEPSYOKSLQTYLKSSG-----SVASLPDS 326
 QY 558 RRTYSRIGRTGTLOKPPSMRSISAVDTAAORAKVAYOYTLVSRQKPLGPNNSNQAS 617
 DB 327 DRSSS---SSQESLNRPFSSSKMSTANPSTVA---GRVSA-----GSSKVAS 365
 QY 618 MTGAAGSSSLPRPR---LINSNGGTPATTPGSPVTPRRGAGVQSQSPGSRSTSPSTKL 673
 DB 366 ---SLPESLORSREDIDVNAAGAKAHAAAGOSV---RSGRIGAGALNAGSYA---SLGK 416
 QY 674 RDQYG 678
 DB 417 RTDYG 421

RESULT 14
 ABB81781
 ID ABB81781 standard; protein; 432 AA.
 XX
 XX ABB81781;
 AC
 XX
 DT 24-SEP-2002 (first entry)
 DE
 DE Proteasome subunit nu 47.52.
 XX
 KM Proteasome subunit nu 47.52; protein metabolic disorder.
 XX
 OS Unidentified.
 OS
 PN CN139447-A.
 XX
 PD 13-MAR-2002.
 XX
 PF 21-AUG-2000; 2000CN-00119655.
 XX
 PR 21-AUG-2000; 2000CN-00119655.
 PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI; 2002-455349/49.
 DR N-PSDB; ABB83145.
 XX
 PT New polypeptide-proteasome subunit nu 47.52 and polynucleotide for
 PT encoding such polypeptide.
 PS
 PS Claim 1; Page 26-27 (Disclosure); 33pp; Chinese.
 CC
 CC The sequence represents the novel proteasome subunit nu 47.52 of the
 CC invention. The invention relates to the proteasome subunit nu
 CC 47.52, and the polynucleotides encoding it. The present invention also
 CC discloses the method of applying the polypeptide in treating various
 CC diseases, such as protein metabolic disorder. The antagonist of the
 CC polypeptide and its treatment effect are also disclosed
 CC
 XX
 SQ Sequence 432 AA:
 Query Match 6.5%; Score 497.5; DB 5; Length 432;
 Best Local Similarity 32.7%; Pred. No. 2.3e-28;
 Matches 136; Conservative 75; Mismatches 154; Indels 51; Gaps 11;
 QY 237 SALK-----NTNNGVGLDEADNIGLERPPTRMIRPLHSAVSSSLRPKNVN--DVTGDAG 291
 DB 31 SAKFVPAKPTSGN-----PANSARKP-----GSAGPKVGASKEGAG 70
 QY 292 AYMTESESEFSEVVPOLNTHADMDIYKQVLIIISDKNADMEKRDALKKIRALLI-- 349
 DB 71 TVBEDDFIKAFITVPSIQIYSRELEBTLNKIREIISDDKGDMDQDANALKKIRSLVAG 130
 QY 350 -LSYHTQPOFVAVOLKELISFVDILKEELRSQYIREACTITIAVMSKTLNKLDAFCMSI 408
 DB 131 AAGYDCFFQHLRLIDGALKLS-----AKDIRSQVIREACTITIAVMSKTLNKLDAFCMSI 185
 QY 409 LEHNLINLIONSASVIVASASTIAKYLIIKYTHAPRLKIYDTLINSKSKDIRSTLCELMV 468
 DB 186 VPTLFLINVPNSAKVMAVMSGCAALFIRHTRHVPPLILITNSCT--SKSVFVRRRSFEFLD 244
 QY 469 LTFEEMOTKALERNATVLRITLKKSIGDADCDARHRHRYVMAFRHRFPPELLADQIYGTLD 528
 DB 245 LILQEMOETHLERHAALVETIKKGIDDAARVREARKTYMGLRNHFPGEAETLYNSLE 304
 QY 529 IAAQRALEREREGGGGTGTGTAPEYTRTYSRIGRTGTLOKPPSMRSISAVDTAA 588
 DB 305 PSYOKSLQTYLKSSG-----SVASLPDSRSSS---SSQESLNRPFSSSKMSTANPSTVA 355
 QY 589 ---AORAKVAYOYTLVSRQKPLGPNNSNQASWTGA---SGSLPRPLINSNS 636

Db 356 GRVSAGSSKASLPGLSRSRSDIGVNAAGAKAHHAAGOSVRSGLGAGALNAGS 411

RESULT 15
ADJ68457
ID ADJ68457 standard; protein: 1972 AA.

XX AC ADJ68457;
XX DT 06-MAY-2004 (first entry)
XX DE Human heart mitochondrial protein as a therapeutic target SeqID263.

XX KM mitochondrial; human; screening assay; diabetes mellitus;
XX KM Huntington's disease; osteoarthritis;
XX KM Leber's hereditary optic neuropathy; LHON;
XX KM mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
XX KM myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
XX KM neuroprotective; neurotropic; antidiabetic; anticonvulsant; antiarthritic;
XX KM osteopathic; ophthalmological; cytoskeletal.

XX OS Homo sapiens.
XX PN NO2003087768-A2.
XX PD 23-OCT-2003;
XX PF 04-APR-2003; 2003MO-US010870.
XX PR 12-APR-2002; 2002US-0372843P.
XX PR 17-JUN-2002; 2002US-038987P.
XX PR 20-SEP-2002; 2002US-0412418P.

XX PA (MITO-) BUCK INST AGE RES.
XX PI Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX PI Warnock DE;
XX DR WPI; 2003-845369/78.

XX PT Identifying a mitochondrial target for drug screening assays and for
XX PT treating diseases associated with altered mitochondrial function,
XX PT comprises detecting a modified polypeptide in a sample and correlating
XX PT with the disease.

XX PS Claim 1; SEQ ID NO 263; 180pp; English.

XX CC This invention relates to novel mitochondrial targets that can be used
XX CC for therapeutic intervention in treating a disease associated with
XX CC altered mitochondrial function. Specifically, it refers to a method for
XX CC identifying proteins of the human heart mitochondrial proteome that are
XX CC useful for drug screening assays, as well as therapeutic targets. The
XX CC present invention describes a method for identifying such proteins that
XX CC can be used in the treatment of various diseases associated with altered
XX CC mitochondrial function including diabetes mellitus, Huntington's disease,
XX CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX CC compositions have neuroprotective, neurotropic, antidiabetic, and
XX CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX CC cytoskeletal activities. This polypeptide sequence is a human heart
XX CC mitochondrial protein of the invention.

XX SO Sequence 1972 AA;

Query Match 4.3%; Score 328; DB 7; Length 1972;
Best Local Similarity 18.7%; Pred. No. 1.7e-14;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

QY 61 SHFPIAKSLAEFLIKRLGSDPNATATVLPVIVIRLGDSDRYREKAQLLRDMEH 120
DB 653 TNQVWQMKLHYA-LIAQKGNFSKTSIAQVVLGDLVVKIDVK--CQNAKEMTAIAEA 709

QY 121 RVLPQALIDKLATSCFKRKNKAVEEFLQTI VNALEHYGTQQLSVRYVTPVYCALLGDP 180
DB 710 CMLPMTA--EGVSNMAFSQKPNKNOSETLNLWLSNAIKERGFGLNVKAPISVYKTAIAAT 767

QY 181 TYNVREAAIQTVLELYKVVGDRLRPDLRMM--DDVPASLAMLBEQFDQVKGGLLPSA 238
DB 768 NPAYRTAATLTLGVMLTVG---PSLKNFFEDKPA-LLSQIDAEFB--KQGGOSPPAP 820

QY 239 LKNTNGNV-GLDEADNIGLRERPTRMIKRPLSAVSSSLRPKNVNDVTGAGAVTMS 297
DB 821 TRGISKHSSTGDBED-----GDEBDGSDNDV----- 849

QY 298 FESSFEVPEQLNIFPAKMDIYKQVLTISPKADMEKRVDAKKIRALLSYHTQPO 357
DB 850 ----DLFRTETI-----SDKITSELVSGIKGN--WKIRKGLDEVAQII-----NDAK 892

QY 358 FAVAVLKEIUSLSPVILKEELBSQVIREACITIAVMSKTLRKLDAPFCSILEHJINLQ 417
DB 893 FIQPNIGELPTA---LKGRLNDS-----NKT-----LVQQTINLQ 925

QY 418 NSAKVIASASTIALKYIIKYTHAPLLKLYTDTLNSKSKDIRSTLCELMVLFPBWQK 477
DB 926 Q---LAVMGPNIKQHVNLGIP-IITVLDGSKXNVAAALATV----- 965

QY 478 ALERNATVIRDTLKKSIGDADCDARRHSRYAVAFRRHPELADQIYGTLDIAQRALE 537
DB 966 ---NAMAEGTGMKEWLEBEDLSEB-----LKKENPFILROELGWL---AEK--- 1005

QY 538 EREGGGGGTGTCTGTAPETRTVSRIGRTPTLQKTPMSMSISAVDPAANOR---AKV 594
DB 1006 ----LPTLRST-----PDLILCVPHLISCLEDRGVDVK 1036

QY 595 RAQYTL-----YSROKPLG--PNNNSQASMTGAASGLSPRPINSNGGTPAT 643
DB 1037 KAQDALPFPMMHLGTEKAKATGKXKLPSTKD-----VLAMEKAKYNNAPAKAP--- 1086

QY 644 PGSVTPPRGRAGVSQSGPGRST---SPSTKLDQYGIGNYRGATGALPKAAGIP 699
DB 1087 PTKATSKPWGSGAPAKFPASAPADCISSSTPEKPD-----PKKA--- 1127

QY 700 RSTASRSTSPRSGGGLMKRMYSTGASRTPPRNNVPRSPARLILAGSREAHITLG 759
DB 1128 ----KAPGSSKAKSAOG--KMPKSTK-----LKEDS----- 1154

QY 760 VDDGQPDVY---SGDYMSGGMGRGKLM-----GRDE-SDIDSEASVCS---ER 805
DB 1155 --DKSGPIFIVPNKQKQRMKDEK-GLKYLKKNFTTPREYIEQKTMSSCVAAKVLQDE 1211

QY 806 SFDSSVTRGNKSNVLSGSHRTLDMWSTORAPPDIEITIIQFCASHTWSERKGLISLQY 865
DB 1212 MHSDFOHNKA-----LAVVWD-----HLESEKGVIGC--- 1241

QY 866 LADGKELTQQLKCVLDMFRKMPDMTHKVSFLDTVELTVHANE---TSRNGSSSC 922
DB 1242 ----DLTLKMLTLRFPTNLSVLKMALEYLKLTLTSEEYHHTENBASFF 1290

QY 923 LTRLENKLG--TDLNMSHKSIMKTLQVNEHYEFPQLOLKEFLRISDSPTPTKTRIA 980
DB 1291 IYLVVYKGEPRDVIRKQVRAILNRMCLVY-----PASKKPFIMESTKSKSQRAE 1343

QY 981 IIRFLTDLANTYCKSDSPSDSQACERT---VLKLAQLADQKSMELRSQARSCIVALLY 1037
DB 1344 CLEELGLVESY-----GMNVCOPFGKALKEIVHIGDRNVAVNAALNTIVTY 1394

QY 1038 NLNTPQMTLLADLPKVQDSARSCHSHMRQSGSCNSGANSPPSSPPLSSPPLQSP 1097
DB 1395 NVHGQVFLKLNL---SEKMSMLBERIK-----SAKRPSAAPKQVEEKFORAO 1443

QY 1098 SV-----GPFASLQSHHQLSISSTSPRSQSVYQELLFSSSELDIQHNIOKT--- 1145
DB 1444 NISSNANMLRKQPADWMSKLNQARSMSGHPEAAQ--VWRRE--FOLDDEINENDGYVAC 1500

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 14:27:58 ; Search time 12424 Seconds
(without alignments)
19375.365 Million cell updates/sec

Title: US-10-030-850-1
Perfect score: 5145
Sequence: 1 gaattcgagcagcgcgcgc.....ggattaaatcaataaaaa 5145

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hnc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gsa1:*
10: gb_gsa2:*
11: gb_gsa3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	925.2	18.0	931	1	A1513688 GH26741.3
2	836.6	16.3	932	2	BG633422 GH26741.3
3	803.2	15.6	843	10	CNS000ADB
4	762.8	14.8	766	1	A1405248 GH25189.5
5	741.2	14.4	856	2	BG633353 GH25189.3
6	719	14.0	720	2	BF485581 AT19621.5
7	716	13.9	716	7	CK658052 LP14575.5
8	692	13.4	692	2	BF504405 AT05759.5
9	645	12.5	656	1	A1297625 LP11963.5
10	619.6	12.0	642	10	AG977859
11	604.4	11.7	606	2	BG634723
12	597	11.6	626	2	BF493211 AT01524.5
13	586	11.6	615	2	BF491691 AT28490.5
14	581	11.5	640	7	CK657543 LP23927.3
15	587.2	11.4	631	2	BF4917453 RL13234.5
16	580	11.3	580	2	BF497552 AT11747.5
17	574.2	11.2	579	2	BF502300 AT17981.5
18	552	10.7	576	10	AG951156
19	551	10.7	551	1	A1543213
20	543	10.6	566	1	A1294703
21	530.6	10.3	559	1	A1945450
22	517.8	10.1	523	1	A1945833

23	509.4	9.9	511	7	CO314038	CO314038 EK234750.
24	507.6	9.9	561	10	AG973742	AG973742 Drosophila
25	500	9.7	500	1	A1456404	A1456404 LD36419.5
26	486.6	9.5	629	10	AG912861	AG912861 Drosophila
27	476.2	9.3	512	2	BF492167	BF492167 AT29056.5
28	473.4	9.2	475	7	CO307834	CO307834 EK25581.8
29	471	9.2	494	7	CO305613	CO305613 EK24910.7
30	469	9.1	518	2	BF486127	BF486127 AT20323.3
31	468	8.9	484	10	AG962276	AG962276 Drosophila
32	450	8.7	450	7	CO284242	CO284242 EK16555.8
33	445.6	8.7	861	10	CNS01100	AL1100530 Drosophila
34	435	8.5	498	2	BG639454	BG639454 LP11963.3
35	425.4	8.3	427	7	CO289204	CO289204 EK07044.0
36	420	8.2	1101	10	CNS014X8	AL104726 Drosophila
37	399.2	7.8	437	1	A1946614	A1946614 bs28905.Y
38	394.8	7.4	412	7	CO270671	CO315184 EK238613.
39	382.6	7.4	572	1	AA541072	CO270671 EK09974.5
40	379	7.0	403	1	AM943097	AA541072 LD20859.5
41	360	7.0	359	7	CO263714	AM943097 LD31673.3
42	359	7.0	303	7	CO286919	CO263714 EK089357.
43	298.2	5.8	285	7	CO268707	CO286919 EK064252.
44	279.2	5.4	285	7	CO268707	CO268707 EK092448.
45	274	5.3	460	2	BE977670	BE977670 bs67e10.Y

ALIGNMENTS

RESULT 1
LOCUS A1513688 931 bp mRNA linear EST 23-APR-2001
DEFINITION GH26741.3prime GH Drosophila melanogaster head pot2 Drosophila
melanogaster cDNA clone GH26741 5, mRNA sequence.
A1513688
A1513688.2 GI:13759214

ACCESSION A1513688
VERSION A1513688
KEYWORDS EST
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 931)
Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)

TITLE On Mar 16, 1999 this sequence version replaced gi:4417834.
JOURNAL Other ESTs: GH26741.3prime
COMMENT Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estefruitfly.berkeley.edu
hit genomic AE005593; arm:3L (2079555,2112669)
estimated-cyto:78A2-78C7: 04/10/2001
Plate: GH.267 row: D column: 5
High quality sequence stop: 616
POLYA=No.

FEATURES
source location/Qualifiers
1..931
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH26741"
/sex="male and female"
/dev stage="adult"
/lab_host="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: POT2; Site: 1; EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
POT2. Plasmid cDNA library."

ORIGIN

DEFINITION	GH26741.3prime GH Drosophila melanogaster head POT2 Drosophila melanogaster cDNA clone GH26741 3, mRNA sequence.
ACCESSION	B6633422
VERSION	B6633422.1
GI	GI:13759213

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Enhydryoidea: Drosophilidae: Drosophila.

[illegible]

Lewis, S. and Rubin, G.M.

JOURNAL
Unpublished (2001)

COMMENT
CUIR1_ES18: 0828/41.0921me
Contact: Stan]eton M

BDGP
Biology Department

One Cyclotron Rd, Berkeley, CA 94720, USA

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu

more T residues at the beginning of the sequence, this clone was

polyadenylated. The resulting poly-T sequence has been removed

estimated-cyto:78A2-78C7: 04/10/2001

High quality sequence stop: 799.

FEATURES	LOCATION/QUALITY
SOURCE	1 932

```
/organism="Drosophila melanogaster"
```

```
/db_xref="taxon:7227"
```

```
/sex="male and female"
```

```
/dev_blage=auuc  
/lab_host="PH5 - a]nha"
```

```
/clone_lib="GH Drosophila melanogaster head POT2"
```

XhoI; sized fractionated cDNAs were directly ligated

ORIGIN	DATE	DESCRIPTION	AMOUNT	CHECK NO.	ACCOUNT NO.
...

16 38: Score 836 6: DB 2: Tienath 932:

Best Local Similarity 98.7%; Pred. No. 1.1e-233;

DD 332 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 12

4343 GCAATCGATGGTTCGCAAGCTGCGGTCTTCTGCAATCGTCAAGCTGTACTTGTTCCTGGC

Db 872 GCAATCGATGGTTCGCAAGCGCGTCTTCTGCATCGTAAAGCTGACTTTGTTCTGCGC

4403 CGAGGAGAGGTCAAGCCGAAGCTGTCAAGTCTAATCCAGCAAGTTAGGCTCTCA

812 AGAGGAGAGTCAAGCCGAGTCTAATCCAGCAAGTTAGGCTCTCA/

[illegible]

97

4323 CCCCCCHCVCICACIACA ICCCCAACCC - CCCCCTT

Db . 692 CTCGGCGCATCGTCATGATTGCGGGAGCCCTTATAGGATTCCCTGCAICAGCACCA

4581 CCACAACAGACACAGACGGGTTGCTTCCCTCGGCTGAGAAGGAAGTGAGGAGCGGC

Db 632 CCACAAACAAGACACGACGGCTTGGCTTCCCTCGGCTGAGAAGGAATGAGGAGCGGC

4641 GGACATTAAATAATATTATTATTAACCTAATTATTATTAACCTAATTAATTACCGATCA

RESULT 4
 A1405248 766 bp mRNA linear EST 23-APR-2001
 LOCUS GH25189.5prime GH Drosophila melanogaster head pot2 Drosophila
 DEFINITION melanogaster cDNA clone GH25189 5, mRNA sequence.
 ACCESSION A1405248 GI:13759105
 VERSION A1405248.2 GI:13759105
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 766)
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S. and Rubin, G. M.
 BDGP/HMI Drosophila EST Project
 TITLE Unpublished (2001)
 JOURNAL On Feb 8, 1999 this sequence version replaced gi:4248335.
 COMMENT Other ESTs: GH25189.3prime
 Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 hit genomic AB003593: arm:3L [20799555,21112669]
 estimated-cyto:78A2-78C7: 04/10/2001
 Plate: GH.251 row: H column: 5
 High quality sequence stop: 634
 POLYA=No.

FEATURES
 source
 1..766
 /location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="GH25189"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DHS - alpha"
 /clone_1lb="GH Drosophila melanogaster head pot2"
 /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pot2. Plasmid cDNA library."

ORIGIN
 Query Match 14.8%; Score 762.8; DB 1; Length 766;
 Best Local Similarity 99.7%; Pred. No. 5.1e-212;
 Matches 764; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1257 AGAATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 1316
 |||||
 1 AGAATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 60
 |||||

1317 AACAGCGGAGAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 1376
 |||||
 61 AACAGCGGAGAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 120
 |||||

1377 ACACATGACCAAGCTGCTCAAGATCTACAGACACTCTGAATCAATCAAGTCGAG 1436
 |||||
 121 ACACATGACCAAGCTGCTCAAGATCTACAGACACTCTGAATCAATCAAGTCGAG 180
 |||||

1437 GACATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 1496
 |||||
 181 GACATTAAGTATGATGCTTCTGCTGAGACATTTTGAGACACCTGATTAATTATACG 240
 |||||

1497 GCGTTGAAAAGAAATGCCACCGTACTAGAGACACCTTAAATAATTCATTGGCGATGA 1556
 |||||
 241 GCGTTGAAAAGAAATGCCACCGTACTAGAGACACCTTAAATAATTCATTGGCGATGA 300
 |||||

1557 GACATGATGACACCGCGCATTCAGATACGGCTATTGGGCTTTCAGAGCGTCACTTCCA 1616
 |||||
 301 GACATGATGACACCGCGCATTCAGATACGGCTATTGGGCTTTCAGAGCGTCACTTCCA 360
 |||||

RESULT 5
 B6633353/ 856 bp mRNA linear EST 23-APR-2001
 LOCUS GH25189.3prime GH Drosophila melanogaster head pot2 Drosophila
 DEFINITION melanogaster cDNA clone GH25189 3, mRNA sequence.
 ACCESSION B6633353 GI:13759104
 VERSION B6633353.1 GI:13759104
 KEYWORDS EST.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 856)
 Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
 Lewis, S. and Rubin, G. M.
 BDGP/HMI Drosophila EST Project
 TITLE Unpublished (2001)
 JOURNAL Other ESTs: GH25189.5prime
 COMMENT Contact: Stapleton, M.
 BDGP
 Lawrence Berkeley National Lab
 One Cyclotron Rd, Berkeley, CA 94720, USA
 Fax: 510 486 6798
 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
 Based upon the presence of a XhoI site followed by a run of 14 or
 more T residues at the beginning of the sequence, this clone was
 polyadenylated. The resulting Poly-T sequence has been removed. hit
 genomic AB003593: arm:3L [20799555,21112669]
 estimated-cyto:78A2-78C7: 04/10/2001
 Plate: GH.251 row: H column: 5
 High quality sequence stop: 363.
 POLYA=No.

FEATURES
 source
 1..856
 /location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /clone="GH25189"
 /sex="male and female"
 /dev_stage="adult"
 /lab_host="DHS - alpha"
 /clone_1lb="GH Drosophila melanogaster head pot2"
 /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2:
 XhoI; Sized fractionated cDNAs were directly ligated into
 pot2. Plasmid cDNA library."

ORIGIN
 Query Match 14.8%; Score 762.8; DB 1; Length 766;
 Best Local Similarity 99.7%; Pred. No. 5.1e-212;
 Matches 764; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1617 GAGCTGGCGGATCAATATATAGAACATTAAGATAGCTGCCAGCGGATTAAGAAAG 1676
 |||||
 361 GAGCTGGCGGATCAATATATAGAACATTAAGATAGCTGCCAGCGGATTAAGAAAG 420
 |||||

1677 GAACGAGAGGGCGCGGAGAGAGAGAACTGTACTGGGACTTGGGACTTGCACCTGAAGC 1736
 |||||
 421 GAACGAGAGGGCGCGGAGAGAGAACTGTACTGGGACTTGGGACTTGCACCTGAAGC 480
 |||||

1737 AGACGACTGTATCCCGCATTTGGCGGAGACACCTGGAACCTTGCAAAAGCCAGCGCTAGT 1796
 |||||
 481 AGACGACTGTATCCCGCATTTGGCGGAGACACCTGGAACCTTGCAAAAGCCAGCGCTAGT 540
 |||||

1797 ATGAGATCCATTTTACGCGGTGACACTCGCGCTCTCAACGAGCCAAAGTTAGAGCGCA 1856
 |||||
 541 ATGAGATCCATTTTACGCGGTGACACTCGCGCTCTCAACGAGCCAAAGTTAGAGCGCA 600
 |||||

1857 TATACACTATATTTCCAGGCAAAAGAAACCTTTAGACCTATATTTCCAAACGAGCATCG 1916
 |||||
 601 TATACACTATATTTCCAGGCAAAAGAAACCTTTAGACCTATATTTCCAAACGAGCATCG 660
 |||||

1917 ATGACAGAGAGAGAGAGAGAGATGATCACTGCGCCAGACCTCGCTGAATTCATAGCGGT 1976
 |||||
 661 ATGACAGAGAGAGAGAGAGATGATCACTGCGCCAGACCTCGCTGAATTCATAGCGGT 720
 |||||

1977 GGCAACCAAGCTACCAACCGCGGATCGTTACACCAACCGCGCGCGG 2022
 |||||
 721 GGCAACCAAGCTACCAACCGCGGATCGTTACACCAACCGCGCGCGG 766
 |||||

Accession	Sequence	Position
Dd	ATGGAGAGGAAAGTGGTGGTGGCTTGACGCTTAACAAGATCAACCCGTCGACGTTG	480
Qy	CTTCAGAGCTCACTTACTTAACCTGGGGATTGSCATCAAGGGGGAACATGAGAGCTGCC	3932
Dd	CTTCAGAGCTCACTTACTTAACCTGGGGATTGSCATCAAGGGGGAACATGAGAGCTGCC	540
Qy	AATAAGCACTTTAGTCGATCGATCGCAATGCTGCTTAACTTTGAGAGCGGAGCATACG	3992
Dd	AATAAGCACTTTAGTCGATCGATCGCAATGCTGCTTAACTTTGAGAGCGGAGCATACG	600
Qy	GACGTGTCATCGCTGGGCTGCACATGCTCACTAAGATTATGAGAGCAACAAATGCGT	4052
Dd	GACGTGTCATCGCTGGGCTGCACATGCTCACTAAGATTATGAGAGCAACAAATGCGT	660
Qy	CACAACTGGATGCACTTTCTAGAGCTGATTTGCTGGAAGATCATCCATGCTATCAAC	4112
Dd	CACAACTGGATGCACTTTCTAGAGCTGATTTGCTGGAAGATCATCCATGCTATCAAC	720

RESULT 7	LOCUS	DEFINITION
CK658052	CK658052	716 bp mRNA linear EST 30-JAN-2004 LP14575.5: rine LP <i>Drosophila melanogaster</i> larval-early pupal P0T2 <i>Drosophila melanogaster</i> cDNA clone LP14575 5, mRNA sequence.

ACCESSION	CK658052	
VERSION	CK658052.1	GI:41401577
KEYWORDS	EST.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	

REFERENCE	1 (bases 1 to 716)
AUTHORS	Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G. I.,
TITLE	Lewis, S., and Rubin, G. M.
JOURNAL	BGP/HMI Drosophila EST Project
COMMENT	Unpublished (2001) Contact: Stapleton, M.

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Plate: LP145 row: G column: 3
High quality sequence stop: 562.

FEATURES	Location/Qualifiers
source	1. . 716

```

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="LPI4575"
/sex="male and female"
/dev_stage="larvae-pupae"
/lb_host="DH5-alpha"
/clone_lib="LP Drosophila melanogaster larval-early pupal
pot2"
/note="Organ: whole body; Vector: pot2; Site 1: EcoRI;
Site 2: XhoI. Sized fractionated cDNAs were directly
ligated into pot2. Plasmid cDNA library."

```

ORIGIN

Query Match	13.9%	Score 716	DB 7	Length 716
Best Local Similarity	100.0%	Pred. No.	3e-198	
Matches 716	Conservative	0	Mismatches	0
			Indels	0
			Gaps	0

QY 881 CGTTTCGATCACGCGCCAAAAACCAATGTGAACAATGTGACCGGTGAATGCCGGGC 940

Db 1 CGTTTCGATCACGCGCCAAAAACCAATGTGAACAATGTGACCGGTGAATGCCGGGC 60

QY 941 CGTAAACATGGAATTTTCGAATTCAGCTTGAAGTGTGCCCCGAATGGAACATTTTCCA 1000

Db	61	CGTACACATGGAAATCTTTGGAACTCTAGCTTTGGAGTGTGCCCGCAATTGAACATCTTCCA	120
QY	1001	CGCTTAAGACATGACGATATCTTACAAAGCAAGTACTGTATCATCATGTATATAAACCGC	1060
Db	121	CGCTTAAGACATGACGATATCTTACAAAGCAAGTACTGTATCATCATGTATATAAACCGC	180
QY	1061	AGACTGGAGAAAAGTGTGATCTCTCAAGAAGATCAGGGCAATTGGTCTCATTTCTCAGCTA	1120
Db	181	AGACTGGAGAAAAGTGTGATCTCTCAAGAAGATCAGGGCAATTGGTCTCATTTCTCAGCTA	240
QY	1121	TCACACTCAGCCGACGTTTGTGCTGTACAGCTAAAGAAATGTCTTTAAGCTTCTGTGA	1180
Db	241	TCACACTCAGCCGACGTTTGTGCTGTACAGCTAAAGAAATGTCTTTAAGCTTCTGTGA	300
QY	1181	CATCTCTAAGAGAAATTACGATTCACAGGTGATCCGGAGAGCGTGCATCACATCCGCTA	1240
Db	301	CATCTCTAAGAGAAATTACGATTCACAGGTGATCCGGAGAGCGTGCATCACATCCGCTA	360
QY	1241	CATGCTTAAGACGCTGAGAAATTAACATAGATGCTTCTGTGTGAGACATTTTGGAGCACT	1300
Db	361	CATGCTTAAGACGCTGAGAAATTAACATAGATGCTTCTGTGTGAGACATTTTGGAGCACT	420
QY	1301	GATTAAATTTAATACAGAACAGCGCGAAGGTCAATTGCATCCGCTTCCACAAATAGCTGTGA	1360
Db	421	GATTAAATTTAATACAGAACAGCGCGAAGGTCAATTGCATCCGCTTCCACAAATAGCTGTGA	480
QY	1361	GTATATCATTTAAGTATTCACATGTCACCAAGCGTGTCCAAATCTACACAAACCTCGAA	1420
Db	481	GTATATCATTTAAGTATTCACATGTCACCAAGCGTGTCCAAATCTACACAAACCTCGAA	540
QY	1421	TCAATCAAAGTCCGAAGGACATAAAGTCCACACTGTGTGAGCTGATGGTGTGCTCTTGA	1480
Db	541	TCAATCAAAGTCCGAAGGACATAAAGTCCACACTGTGTGAGCTGATGGTGTGCTCTTGA	600
QY	1481	GGAGTGCACAGCAAGGCGTTTGGAAAAGAAATGCCACCTGACTTAAAGGACACTTTAAAAAA	1540
Db	601	GGAGTGCACAGCAAGGCGTTTGGAAAAGAAATGCCACCTGACTTAAAGGACACTTTAAAAAA	660
QY	1541	ATCATTTGGGATGACGACCTGGCATGACGCGCGCAATTCAGATPAGCCCTAATTTGGG	1596
Db	661	ATCATTTGGGATGACGACCTGGCATGACGCGCGCAATTCAGATPAGCCCTAATTTGGG	716

RESULT 8					
BF504405	BF504405	692 bp	mRNA	linear	EST 19-APR-2001
LOCUS					
DEFINITION	AT05759.5	prime AT	<i>Drosophila melanogaster</i>	adult testes	POTB7
	<i>Drosophila melanogaster</i>	cDNA clone	AT05759 5,	mRNA	sequence.

ACCESSION	BF504405	
VERSION	BF504405.2	GI:13687029
KEYWORDS	EST.	
SOURCE	Drosophila melanogaster (fruit fly)	
ORGANISM	Drosophila melanogaster	

REFERENCE
AUTHORS

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidei; Drosophilidae; Drosophila.

1 (baes 1 to 632)

Stapleton, M., Bröckstein, P., Hong, L., Agbayani, A., Baxter, E.,

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 692)	Shapleton, M., Brocksheini, P., Hong, L., Agbawani, A., Baxter, R.			
	Berman, B., Carlson, J., Champagne, M., Chavez, C., Chew, M., Dorr, M.,			
	Fairlin, D., Frise, E., George, R., Gonzalez, M., Guatin, H., H., H.,			
	P., Liao, G., Miranda, A., Mista, S., Mungall, C. J., Nuno,			
	Pacble, S. J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,			
	Lewis, S. E., Ceiniker, S. and Rubin, G. M.			
	BDGP/HMMI AT Drosophila Est Project			
	unpublished (2000)			
	On Dec 6, 2000 this sequence version replaced gi:11587706			

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_esc@fruitfly.berkeley.edu
hit genomic AE003593: arm.3L [20799555,2112669]

OY	891	GGGTGAGAGGACGACCCACGAGATGATTAAAGCGGACCACTACATCGGCGGTTTGCGCA	890
Db	301	GGGTTGAGAGGACCGACCCACGAGATGATTAAAGCGGACCTACATCGGCGGTTTGCGCA	360
OY	891	TCACGTGCGCCCAAAACCCAAATGTGACATGTGACCGGTGATGCGCGGCGCGTAACCAATG	950
Db	361	TCACGTGCGCCCAAAACCCAAATGTGACATGTGACCGGTGATGCGCGGCGCGTAACCAATG	420
OY	951	GAATCTTTGGAATCTAGCTTTGAGGTGTCGCCGCAATTGAACATCTTCCAGCTTAAGGAC	1010
Db	421	GAATC-TTCGAATCTAGCTTTGAGGTGTCGCCGCAATTGAACATCTTCCAGCTTAAGGAC	479
OY	1011	ATGGAACCATATCTAAAGCAAGTACTAGTGAATCATCACTGATATAAAAACGCGACCTGGAG	1070
Db	480	ATGGAACCATATCTAAAGCAAGTACTAGTGAATCATCACTGATATAAAAACGCGACCTGGAG	539
OY	1071	AAACGTGAGATGCTCTCAAGAAGATCAAGGCGATGCTCATTCCTAGCTATCAACCTGAG	1130
Db	540	AAACGTGAGATGCTCTCAAGAAGATCAAGGCGATGCTCATTCCTAGCTATCAACCTGAG	599
OY	1131	CCGCGAGTTTGCCTGTACAGCTAAAGGAATTGCTTAAAGCTTGCCTGTGACATCTCC	1187
Db	600	CCGCGAGTTTGCCTGTACAGCTAAAGGAATTGCTTAAAGCTTGCCTGTGACATCTCC	656

RESULT 10			
AG977859/c			
LOCUS	AG977859	642 bp	DNA linear
DEFINITION	Drosophila simulans DNA, clone: DSM-02J17.F.fa, genomic survey		

ACCESSION	AC977859
VERSION	AC977859.1
KEYWORDS	GS5
SOURCE	<i>Drosophila simulans</i>
ORGANISM	<i>Drosophila simulans</i>

REFERENCE
AUTHORS
1
Hatтори, M., Toyoda, A., Murakami, K., Kuroki, Y., Fujiyama, A.,

TITLE BAC end sequences of library DNBI
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 642)
AUTHORS Hattori, M.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-2005) Maehara Hattori, The Institute of Physical

COMMENT
Clones are derived from the BAC library DSM1
For BAC library availability, please contact Masa-Toshi Yamamoto

Submitted (30-11-2004) by Masahira Hattori,
Riken, Genomic Sciences Center (GSC);
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@gsc.riken.jp; Tel: 81-45-503-9111,
Fax: 81-45-503-9170)
This work was done in collaboration with Yamamoto, M.T. Drosophila
Genetic Resource Center
Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
Tel: 81-75-873-2660 FAX: 81-75-861-0881
PRIMERS
Sequencing : F
LIBRARY

```

FEATURES
  source      location/Qualifiers
1. .642      /organism="Drosophila simulans"
/mol_type="genomic DNA"

```

```

/db xref="taxon:7240"
/clone="DSM1-021J17.F.fa"
/clone_lib="DSM1 Drosophila BAC library"

```

Query Match	12.0%;	Score 619.6;	DB 10;	Length 642;
Best Local Similarity	97.8%;	Pred. No. 6.7e-170;		
Matches 628; Conservative	0;	Mismatches 14;	Indels 0;	Gaps 0;

OY	2040	CAGTCACAGCGAGATGGAGATCCACTCCGCAAGCAAACTGAGGGATCAGTACGGT	2099
Db	642	CAGTCACAGCGAGATCAAGATCCACTCCGCAAGCAAACTGAGGGATCAGTACGGT	583
OY	2100	GGTATTTGGTAAATTCTATCTGTGTGGGCCCATCTGGCCCATACCCCAAAAAGGCTCTCCGAAATA	2159
Db	582	GGTATTTGGTAAATTCTATCTGTGTGGGCCCATCTGGCCCATACCCCAAAAAGGCTCTCCGCAATA	523
OY	2160	CCCCGAAGCACAGCGCACTCTAAGGGAAACCAAGTCCACTAAGGTCAAGGTGTGTGGCTTGA TG	2219
Db	522	CCCCGAAGCACAGCGCACTCTAAGGGAAACCAAGTCCACTAAGGTCAAGGTGTGTGGCTTGA TG	463
OY	2220	AAACGACGATATGTACTCTACAGGTGGGGGGCTTCGACGTACGCGCCGAGAGAAACAACCA	2279
Db	462	AAACGACGATATGTACTCTACAGGTGGGGGGCTTCGACGTACGCGCCGAGAGAAACAACCA	403
OY	2280	GTAAGACCATTCGCGCGCGGCAAGCATCTGTGGCGCAATCCCGTGAAGCAAAATATCATTTA	2339
Db	402	GTAAGACCATTCGCGCGCGGCAAGCATCTGTGGCGCAATCCCGTGAAGCAAAATATCATTTA	343
OY	2340	GGCGTCGAGATGTATGTGACAGCCGACGTATAGTTTCGGGGGACTATCGCGAGCGGCGGA	2399
Db	342	GGCGTCGAGATGTATGTGACAGCCGACGTATAGTTTCGGGGGACTATCGCGAGCGGCGGA	283
OY	2400	ATCGCGATGGGTGAAGGAGCTCATGGGACGATGAGTCTGTATGTGACATCGACTCCGAGGCC	2459
Db	282	ATCGCGATGGGACGAAAGCTCATGGGACGATGAGTCTGTATGTGACATCGACTCCGAGGCC	223
OY	2460	AGTTCTGTGTGTTCAGAAACGATCTTTTGACTCCAGCTACACTAAGGATTAATTAATCGAAC	2519
Db	222	AGTTCTGTGTGTTCAGAAACGATCTTTTGACTCCAGCTACACTAAGGATTAATTAATTCGAAC	163
OY	2520	TACTCATTTAGCGGGAGCCACACCCGCTTGGACTGTGAGACGACGCGGGCCCATTTTGAC	2579
Db	162	TACTCATTTAGTGGGACCCACACCCGCTTGGACTGTGAGACGACGAAAGGGCGCCATTTGAC	103
OY	2580	GACATCGACAGATTAATTCAGTTCTGCGCATCGACGCAATTTGTTGAAGAAGATGCG	2639
Db	102	GACATCGATACGATTAATTCAGTTCTGCGCATCGACGCAATTTGTTGAAAGAAGATGCG	43
OY	2640	CTGATTCAGCCCTCACACGATATCTGGCGGATGGAAAGGAGCTC	2681
Db	42	CTGATTCAGCCCTCACACGATATCTGGCGGATGGAAAGGAGCTC	1

RESULT 11					
BG634723					
LOCUS					
DEFINITION	BG634723	606 bp	mRNA	linear	EST 23-APR-2001
ACCESSION	AT31034.5	prime AT	<i>Drosophila melanogaster</i>	adult testes	POT87
VERSION	<i>Drosophila melanogaster</i>	CDNA clone	AT31034.5	mRNA sequence.	
KEYWORDS	BG634723				
SOURCE	BG634723.1	GI:13762260			
ORGANISM	EST.				
	<i>Drosophila melanogaster</i>	(fruit fly)			
	<i>Drosophila melanogaster</i>				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; <i>Drosophila</i> .				
REFERENCE	1 (bases 1 to 606)				
AUTHORS	Stepleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,				
	Berman, B., Carlson, J., Champ, M., Chavez, C., Chew, M., Dorsett, V.,				
	Fatlan, D., Friese, E., George, R., Gonzalez, M., Guatin, H., Harris, N.,				
	Li, P., Liao, G., Miranda, A., Mista, S., Mungall, C. J., Nuno, U.,				
	Pacleb, Y., Paragás, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,				

TITLE Lewis, S.E., Celniker, S. and Rubin, G.M.
JOURNAL BDGP/HMI At Drosophila EST Project
COMMENT Unpublished (2000)
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: arm:3L [20799555, 21112669]
estimated-cyto:78A2-78C7: 04/09/2001
Plate: AT.310 row: C column: 10
High quality sequence stop: 551.
Location/Qualifiers

FEATURES

source
1..606
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT31034"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_id="AT Drosophila melanogaster adult testes POTB7"
/note="Organ: ADULT testes; Vector: POTB7; Site_1: EcORI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

ORIGIN

Query Match 11.7%; Score 604.4; DB 2; Length 606;
Best Local Similarity 99.8%; Pred. No. 1.9e-165;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

2887 AGATTGGAGACGCTACAGGTGTCCAGCAATATTTCCACGCGAGTTGCACTTAAG 2946
1 AGATTGGAGACGCTACAGGTGTCCAGCAATATTTCCACGCGAGTTGCACTTAAG 60
2947 AACTATTGAATCATATCGAATTTCTACTCAAAGCCCAACCAAGAGCGGCAATGCCA 3006
61 AACTATTGAATCATATCGAATTTCTACTCAAAGCCCAACCAAGAGCGGCAATGCCA 120
3007 TTCTGCGCTCTTACGAGATCTGCGCAATATCACTTAAGACGCGACTTTCCACGCG 3066
121 TTCTGCGCTCTTACGAGATCTGCGCAATATCACTTAAGACGCGACTTTCCACGCG 180
3067 ACCAGAGCGAGGCTCGAGCGGAGCGGCTCTTAAGCTGCGCGGCGGATCGA 3126
181 ACCAGAGCGAGGCTCGAGCGGAGCGGCTCTTAAGCTGCGCGGCGGATCGA 240
3127 AGTCAGAGAGTGGCTGCCAGGCGGAGAGTGGCTTATGAGCTCTGTTAATACCTGAATA 3186
241 AGTCAGAGAGTGGCTGCCAGGCGGAGAGTGGCTTATGAGCTCTGTTAATACCTGAATA 300
3187 CCCCAGCAATGACCTTTTATCTGCGCGAGCTGCGCAAGGTATATCAGAGCTTGGCCGAT 3246
301 CCCCAGCAATGACCTTTTATCTGCGCGAGCTGCGCAAGGTATATCAGAGCTTGGCCGAT 360
3247 CCGTATCCATTCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGC 3306
361 CCGTATCCATTCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGC 420
3307 CTAGTAGCTCTTCATTTAGAGCAGTAGAGTCCCAAGCTTTTGCAGAACTCCCTGAGGAGTC 3366
421 CTAGTAGCTCTTCATTTAGAGCAGTAGAGTCCCAAGCTTTTGCAGAACTCCCTGAGGAGTC 480
3367 CATTTGCTCTGCTTACAGAGCCACCAACCAATTAGCATCAGCTCTACTAGTCCAGCT 3426
481 CATTTGCTCTGCTTACAGAGCCACCAACCAATTAGCATCAGCTCTACTAGTCCAGCT 540
3427 CCGGAGAGTCTTCCGTGAGAGAGAGCTCTTTTCTCGAGCTGAGCATTCAGACACA 3486

Db 541 CCGGAGAGTCTTCCGTGAGAGAGAGCTCTTTTCTCGAGCTGAGCATTCAGACACA 600
QY 3487 ACATTC 3492
Db 601 ACATTC 606

RESULT 12
BF493211 626 bp mRNA linear EST 19-APR-2001
LOCUS AT01524.5prime AT Drosophila melanogaster adult testes POTB7
DEFINITION Drosophila melanogaster cDNA clone AT01524 5, mRNA sequence.
BF493211
ACCESSION BF493211.2 GI:13683797
VERSION
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 626)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Friese, B., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Miera, S., Mungall, C.J., Nunoo, J.,
Pacled, J., Pargava, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S.E., Celniker, S. and Rubin, G.M.
BDGP/HMI At Drosophila EST Project
Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11576512.
Contact: Stapleton, M.

TITLE
JOURNAL BDGP/HMI At Drosophila EST Project
COMMENT Unpublished (2000)
On Dec 6, 2000 this sequence version replaced gi:11576512.
Contact: Stapleton, M.

BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: arm:3L [20799555, 21112669]
estimated-cyto:78A2-78C7: 04/07/2001
Plate: AT.15 row: B column: 12
High quality sequence stop: 500.
Location/Qualifiers

FEATURES

source
1..626
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT01524"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_id="AT Drosophila melanogaster adult testes POTB7"
/note="Organ: ADULT testes; Vector: POTB7; Site_1: EcORI;
Site_2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into POTB7. Plasmid cDNA library."

ORIGIN

Query Match 11.6%; Score 597; DB 2; Length 626;
Best Local Similarity 100.0%; Pred. No. 3e-163;
Matches 597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

14 GCCGCTCCGAATTTGTTGCGGAGAACGAAATGACATTCGAAAGTGGCCCTA 73
30 GCCGCTCCGAATTTGTTGCGGAGAACGAAATGACATTCGAAAGTGGCCCTA 89
74 TCGGAAGCCGAGCACTGATGAGCTTATATCAGAAATGCGCAAGCGGACATGCTGT 133
90 TCGGAAGCCGAGCACTGATGAGCTTATATCAGAAATGCGCAAGCGGACATGCTGT 149
134 GAAGTATAGCTCGCGGAGATCTGTGATCAATTCCTTAGCAGACACAAATCAATGT 193

```

Db      150 GAAGGTACAGCTGCGAGAGATCTGTGATTCCTTAAGGACACAAACCAATTCT 209
Qy      194 CTCACGACGATGGAGATTCTTATTTAGCGTTTATGATGCGTACGCGGAGCCACTT 253
Db      210 CTCACGACGATGGAGATTCTTATTTAGCGTTTATGATGCGTACGCGGAGCCACTT 269
Qy      254 TAAGATTGCAAAAAGTCCCTGGAGGCGTTCTCGAGCTTAATAAGCGATTGGGACGA 313
Db      270 TAAGATTGCAAAAAGTCCCTGGAGGCGTTCTCGAGCTTAATAAGCGATTGGGACGA 329
Qy      314 TTTTATTCATACACGCTTACCGTTCTTCCACATGTCATGATCGCTGGAGACGACG 373
Db      330 TTTTATTCATACACGCTTACCGTTCTTCCACATGTCATGATCGCTGGAGACGACG 389
Qy      374 GGACACAGTCCGAGAGAGCGCAACTTGTGCTGCGGAGCCTCATGAGAGACAGATGCT 433
Db      390 GGACACAGTCCGAGAGAGCGCAACTTGTGCTGCGGAGCCTCATGAGAGACAGATGCT 449
Qy      434 TCGGCCCCAGGCGCTGATCGACAGCTGGCCACTAGCTGTTCCAGACAGAACGCCAA 493
Db      450 TCGGCCCCAGGCGCTGATCGACAGCTGGCCACTAGCTGTTCCAGACAGAACGCCAA 509
Qy      494 GGTGCGCGAGAGATTCTTCAAGCATTTGGAACGCTTTCATGATGAGGACCCAGCA 553
Db      510 GGTGCGCGAGAGATTCTTCAAGCATTTGGAACGCTTTCATGATGAGGACCCAGCA 569
Qy      554 GCTTAGTGTTCGCGCTATATATACACCAAGTTTGACACTTCTCGAGATCCCACT 610
Db      570 GCTTAGTGTTCGCGCTATATATACACCAAGTTTGACACTTCTCGAGATCCCACT 626

```

RESULT 13
LOCUS BF491691 615 bp mRNA linear EST 23-APR-2001
DEFINITION AT28490.5prime AT Drosophila melanogaster adult testes potB7
Drosophila melanogaster cDNA clone AT28490.5, mRNA sequence.

ACCESSION BF491691.2 GI:13756851
VERSION 1.0
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster

REFERENCE 1 (bases 1 to 615)
Stapleton, M., Broksstein, P., Hong, L., Agbayani, A., Baxter, E.,
Bernan, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Miera, S., Mungall, C. J., Nuno, J.,
Pacleb, J., Paragag, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S. E., Celinker, S. and Rubin, G. M.
BDGP/HMNI AT Drosophila EST Project

TITLE Unpublished (2000)
JOURNAL On Dec 6, 2000 this sequence version replaced gi:11574992.
COMMENT Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd. Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AEO03593: atm:3L [20799555, 21112669]
estimated-cyto:78A2-78C7: 04/09/2001
plate: AT.284 row: H column: 6
High quality sequence stop: 596.

FEATURES
Location/Qualifiers

1..615
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT28490"
/sex="male"
/dev_stage="0-3 day old Ore-R males"

/lab_host="Plates AT.10-AT.120: DHS-alpha. Plates
AT.121-AT.319: DHS-alpha Tona"
/clone_1ib="AT Drosophila melanogaster adult testes potB7"
/note="Organ: ADULT testes; Vector: potB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."

Query Match 11.6%; Score 596; DB 2; Length 615;
Best Local Similarity 100.0%; Pred. No. 5.8e-163;
Matches 596; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      14 GCCGTCGGAATTGTTGGCCGAGAGAGAGACATAGCATTTGACGATGGCTCA 73
Db      20 GCCGTCGGAATTGTTGGCCGAGAGAGAGACATAGCATTTGACGATGGCTCA 79
Qy      74 TCGAAGCCACGACGACCTGATGGCTTTATCCAGCAAAATGCCCAAGCCGACATGCTGT 133
Db      80 TCGAAGCCACGACGACCTGATGGCTTTATCCAGCAAAATGCCCAAGCCGACATGCTGT 139
Qy      134 GAAGTACAGCTCCGCGAGAGATCTGTGACATCTCTTTAGCGACACACAACTCAATTGT 193
Db      140 GAAGTACAGCTCCGCGAGAGATCTGTGACATCTCTTTAGCGACACACAACTCAATTGT 199
Qy      194 CTGACGACATNGGATCTCTTATTTGACGTTTATGATGTCATGTCGTCGAGCGGACGACTT 253
Db      200 CTGACGACATNGGATCTCTTATTTGACGTTTATGATGTCATGTCGAGCGGACGACTT 259
Qy      254 TAAGATTGCAAAAAGTCCCTGAGAGCGTTCTCGAGCTTAATAAGCGATTGGGACGCA 313
Db      260 TAAGATTGCAAAAAGTCCCTGAGAGCGTTCTCGAGCTTAATAAGCGATTGGGACGCA 319
Qy      314 TTTTATTCATACACGCTTACCGTTCTTCCACATGTCATGATCGCTGGAGACGACG 373
Db      320 TTTTATTCATACACGCTTACCGTTCTTCCACATGTCATGATCGCTGGAGACGACG 379
Qy      374 GGACACAGTCCGAGAGAGCGCAACTTCTGCTGCGACACTCATGAGACAGAGTGTCT 433
Db      380 GGACACAGTCCGAGAGAGCGCAACTTCTGCTGCGACACTCATGAGACAGAGTGTCT 439
Qy      434 TCGGCCCCAGGCGCTGATCGACAGCTGGCCACTAGCTGTTCAAGACACAAAGGCCAA 493
Db      440 TCGGCCCCAGGCGCTGATCGACAGCTGGCCACTAGCTGTTCAAGACACAAAGGCCAA 499
Qy      494 GGTGCGCGAGAGATTCTTCAAGCATTTGGAACGCTTTCATGATGAGGACCCAGCA 553
Db      500 GGTGCGCGAGAGATTCTTCAAGCATTTGGAACGCTTTCATGATGAGGACCCAGCA 559
Qy      554 GCTTAGTGTTCGCGCTATATATACACCAAGTTTGACACTTCTCGAGATCCCACT 609
Db      560 GCTTAGTGTTCGCGCTATATATACACCAAGTTTGACACTTCTCGAGATCCCACT 615

```

RESULT 14
LOCUS CK657543 640 bp mRNA linear EST 30-JAN-2004
DEFINITION LP23927.3prime LP Drosophila melanogaster larval-early pupal potB2
Drosophila melanogaster cDNA clone LP23927.3, mRNA sequence.
ACCESSION CK657543.1 GI:41401068
VERSION 1.0
KEYWORDS EST.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 640)
AUTHORS Harvey, D., Broksstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,
Lewis, S. and Rubin, G. M.
BDGP/HMNI Drosophila EST Project


```
Db      242 TCCCCAGCGAGTTGACGCTTAAGAACTATTAGATCATATCGGATTCTACTCAAAAGC 301
QY      2983 CAACTACCAAGAGCGCGCATTCGCGCTTCCTTAGGATCTGGCCATATACATACT 3042
Db      302 CAACTACCAAGAGCGCGCATTCGCGCTTCCTTAGGATCTGGCCATATACATACT 361
QY      3043 GTAAAGACAGCACTTCCAGCGACCAAGCCAGGCGCTGCGAGCGGACGGTCTTAAAGC 3102
Db      362 GCAAAAGACAGCACTTCCAGCGACCAAGCCAGGCGCTGCGAGCGGACGGTCTTAAAGC 421
QY      3103 TGGCCAGCTGGGGGGGATCAGAGTCGATGAGCTGGCGCTCCAGGCCAGAGAGCTGCC 3162
Db      422 TGGCCAGCTGGGGGGGATCAGAGTCGATGAGCTGGCGCTCCAGGCCAGAGAGCTGCC 481
QY      3163 TAGTGCCCTGTATAACCTGAATACCCCGCAATGACCCCTTTACTGGCCGACCTGCCAA 3222
Db      482 TAGTGCCCTGTATAACCTGAATACCCCGCAATGACCCCTTTACTGGCCGACCTGCCAA 541
QY      3223 AGGTATATCAGGACTCTGCCGATCCTGCATCCATTGCAATGAGGCGGCAAAAGCCAAA 3282
Db      542 AGGTATATCAGGACTCTGCCGATCCTGCATCCATTGCAATGAGGCGGCAAAAGCCAAA 601
QY      3283 GTTGCAATTGGGGTGGCCAAATTCGCCTTAGTA 3312
Db      602 GTTGCAATTGGGGTGGCCAAATTCGCCTTAGTA 631
```

Search completed: November 30, 2005, 00:40:44
Job time : 12430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 00:05:42 ; Search time 1882 Seconds
(without alignments)
18219.896 Million cell updates/sec

Title: US-10-030-850-1

Perfect score: 5145
Sequence: 1 gaattcgagccgagccgctc.....ggattcaaatcaataaaaa 5145

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_21:*

- 1: geneseqn1980s:*
- 2: geneseqn1990s:*
- 3: geneseqn2000s:*
- 4: geneseqn2001as:*
- 5: geneseqn2001bs:*
- 6: geneseqn2002as:*
- 7: geneseqn2002bs:*
- 8: geneseqn2003as:*
- 9: geneseqn2003bs:*
- 10: geneseqn2003cs:*
- 11: geneseqn2003ds:*
- 12: geneseqn2004as:*
- 13: geneseqn2004bs:*
- 14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5145	100.0	5145	4 AAF24991	Aa124991 Nucleotid
2	149.8	2.9	3153	10 ADB62173	AdB62173 Human CDN
3	149.8	2.9	7888	13 ADB83507	AdB83507 Human mul
4	143.4	2.8	6245	3 AAc64201	Aac64201 Human pol
5	141.8	2.8	2524	10 ADB62363	AdB62363 Human CDN
6	116.6	2.3	1908	10 ADB61959	AdB61959 Human CDN
7	116.6	2.3	5614	10 ADB79874	AdB79874 Rat putat
8	116.6	2.3	5614	14 ADX07257	AdX07257 Cyc1in-de
9	116.6	2.3	6445	4 Aa158443	Aa158443 Human pol
10	116.6	2.3	6445	5 ADQ98654	AdQ98654 DNA encod
11	116.6	2.3	6445	9 ADB48414	AdB48414 Novel hum
12	116.6	2.3	6487	5 Aa158442	Aa158442 Human pol
13	116.6	2.3	6487	9 ADQ98653	AdQ98653 DNA encod
14	116.6	2.3	6487	9 ADB48413	AdB48413 Novel hum
15	116.6	2.3	6491	4 Aa160228	Aa160228 Human pol
16	116.6	2.3	6491	4 Aa160229	Aa160229 Human pol
17	115	2.2	1467	6 Aa160229	Aa160229 Human pol
18	115	2.2	6491	6 Aa160229	Aa160229 Human pol
19	98.4	1.9	6951	10 ADG32852	AdG32852 Human DNA

C	20	92	1.8	608	5 ABV59210	ABV59210 Human pro
C	21	80.2	1.6	695	10 ADF53926	ADf53926 Human con
C	22	52.8	1.0	615	5 ABV59086	ABv59086 Human pro
C	23	52.6	1.0	2000	8 ADA71938	Ada71938 Rice gene
C	24	50.6	1.0	486	5 Aa570766	Aa570766 DNA encod
C	25	49.8	1.0	371	4 AaK63080	AaK63080 Human imm
C	26	49.4	0.9	2000	11 AC137108	Ac137108 Rice stre
C	27	48.4	0.9	2000	11 AC135887	Ac135887 Rice stre
C	28	46	0.9	548	13 AD078622	Ad078622 Novel can
C	29	45.4	0.9	2000	8 ADA71938	Ada71938 Rice gene
C	30	45.2	0.9	561	5 Aa570767	Aa570767 DNA encod
C	31	44.2	0.9	1187	6 ABQ41869	AbQ41869 Oligonuc1
C	32	44.2	0.9	1187	6 ABQ41868	AbQ41868 Oligonuc1
C	33	42.8	0.8	2645	2 AaV62158	AaV62158 HSV-2 str
C	34	42.8	0.8	2811	2 AaV62146	AaV62146 HSV-2 str
C	35	42.8	0.8	3471	2 AaT93650	AaT93650 Herpes ai
C	36	42.8	0.8	3472	2 AaQ48496	AaQ48496 Glycoprot
C	37	42.8	0.8	117213	2 AaV62176	AaV62176 HSV-2 str
C	38	42	0.8	5047	11 ADN95651	Adn95651 Human BEC
C	39	42	0.8	8210	6 ABR31380	AbR31380 Signal tr
C	40	42	0.8	8210	6 ABL70331	Ab170331 Chemicali
C	41	42	0.8	8210	6 Aa561282	Aa561282 Human gen
C	42	41.8	0.8	528	6 ABQ29879	AbQ29879 Oligonuc1
C	43	41.8	0.8	528	6 ABQ29878	AbQ29878 Oligonuc1
C	44	41.8	0.8	1566	13 ADX09689	Adx09689 Plant ful
C	45	41.6	0.8	1883	13 ADX59482	Adx59482 Plant ful

ALIGNMENTS

RESULT 1	AAf24991	standard; cDNA; 5145 BP.
ID	AAf24991	standard; cDNA; 5145 BP.
AC	AAf24991;	
DT	30-APR-2001	(first entry)
DE	Nucleotide sequence of a Drosophila polypeptide designated orbit.	
XX	Orbit; chromosome segregation; mitotic spindle; microtubule; mitosis;	
KW	polyploid cell; microtubule binding; nucleation; motor protein; cancer;	
KM	leukemia cell; solid tumour; ss.	
XX	Drosophila sp.	
OS	Drosophila sp.	
XX	Key	
FT	Location/Qualifiers	
FT	66..4544	
FT	/*tag= a	
FT	/product= "Orbit"	
XX	WO200104295-A1.	
XX	18-JAN-2001.	
XX	11-JUL-2000; 2000WO-GB002662.	
XX	13-JUL-1999; 99GB-00016402.	
XX	24-DEC-1999; 99GB-00030707.	
XX	(Uydu-) UNIV DUNDEE.	
XX	Avides MDC, Deak P, Glover DM;	
XX	MPI; 2001-138345/14.	
XX	P-PSDB; AaB31597.	
XX	Novel Drosophila orbit protein and its human homolog, useful for treating	
XX	tumors and in assays for identifying substances capable of inhibiting	
XX	mitosis.	
XX	Claim 3; Page 84-85; 92pp; English.	

XX The present sequence encodes a Drosophila polypeptide designated orbit.
 CC The orbit protein is required for chromosome segregation. Orbit protein
 CC localises to mitotic spindles and binds microtubules. Orbit mutant
 CC phenotypes result in embryos exhibiting abnormal mitosis and polyploid
 CC cells. Orbit polypeptides are useful for inhibiting mitosis, and for
 CC identifying substances which affect orbit functions such as microtubule
 CC binding, microtubule organizing centre, nucleation activity and
 CC interactions with microtubule motor proteins. Tests for detecting or
 CC sequencing orbit, or its homologue, in a biological sample may be used to
 CC determine orbit sequences within cells in individuals who have or are
 CC suspected to have, an altered orbit gene sequence, for example within
 CC cancer cells including Leukaemia cells and solid tumours such as breast,
 CC ovary, lung, colon etc

Sequence 5145 BP; 1388 A; 1331 C; 1295 G; 1131 T; 0 U; 0 Other;

Query Match 100.0%; Score 5145; DB 4; Length 5145;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 5145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GAATTCGACGACGCGCTCCGAAATTTGTTGCCGAGAACGAGGACGAATGACATTG 60
DB 1 GAATTCGACGACGCGCTCCGAAATTTGTTGCCGAGAACGAGGACGAATGACATTG 60
QY 61 CAACGATGACCTATCGAAGCCGACCTGGATGGCTTTATCCAGCAATGCCCAAG 120
DB 61 CAACGATGACCTATCGAAGCCGACCTGGATGGCTTTATCCAGCAATGCCCAAG 120
QY 121 CGGACATGCGTGTGAAGGTACAGCTGCGGAGGATTTGGTGAATTCCTTAGCGACGAC 180
DB 121 CGGACATGCGTGTGAAGGTACAGCTGCGGAGGATTTGGTGAATTCCTTAGCGACGAC 180
QY 181 CAAATCAATTCCTGCAAGGACATGAGGATTCCTTATGACGGTTGATGCCATGGCTGA 240
DB 181 CAAATCAATTCCTGCAAGGACATGAGGATTCCTTATGACGGTTGATGCCATGGCTGA 240
QY 241 CGGAGACGCACTTTAAGTATGACAAAGTCCCTGAGGCGTTCTCGAGACTAATAAGC 300
DB 241 CGGAGACGCACTTTAAGTATGACAAAGTCCCTGAGGCGTTCTCGAGACTAATAAGC 300
QY 301 GATTGGGCGCATTTTAATGATACACGGCTTACCGTTCGCCAATGATGATCGATCGGC 360
DB 301 GATTGGGCGCATTTTAATGATACACGGCTTACCGTTCGCCAATGATGATCGATCGGC 360
QY 361 TGGGAGACGACGAGGACACAGTCCGAGAGAGCGCACTTCTGCTGCGGACCTCATGG 420
DB 361 TGGGAGACGACGAGGACACAGTCCGAGAGAGCGCACTTCTGCTGCGGACCTCATGG 420
QY 421 AGCAGAGAGTCTTCGCGCCGACGCGCTGATGACAAAGCTGCGCACTAGCTCTTCAAGC 480
DB 421 AGCAGAGAGTCTTCGCGCCGACGCGCTGATGACAAAGCTGCGCACTAGCTCTTCAAGC 480
QY 481 ACAAGAACGCCAAGTGCAGGAGAGTTCCTTCAAGCATTTGTGAACGCTCTTCAAGT 540
DB 481 ACAAGAACGCCAAGTGCAGGAGAGTTCCTTCAAGCATTTGTGAACGCTCTTCAAGT 540
QY 541 ACGGACCCGACGAGCTTAGTGTTCGCGTATATACCAACAGTTTGTGACCTTCCGAG 600
DB 541 ACGGACCCGACGAGCTTAGTGTTCGCGTATATACCAACAGTTTGTGACCTTCCGAG 600
QY 601 ATCCCAAGTTAATGTGAGGAGCGCGCCATCCAACGCTAGTGAATCTACAAGCATG 660
DB 601 ATCCCAAGTTAATGTGAGGAGCGCGCCATCCAACGCTAGTGAATCTACAAGCATG 660
QY 661 TAGGGGATTCGATTTGGCCCAAGCTTCGTCGATGAGCATGTTCTGCTGAAATTTGG 720
DB 661 TAGGGGATTCGATTTGGCCCAAGCTTCGTCGATGAGCATGTTCTGCTGAAATTTGG 720
QY 721 CTATGTTGAGCAAAAGTTGACACAGGTCAACAGAGGGGTCTACTGCTTACGCC 780
DB 721 CTATGTTGAGCAAAAGTTGACACAGGTCAACAGAGGGGTCTACTGCTTACGCC 780

```

```

QY 781 TTAACCAACGATGACGATGAGTGGGCTTGAACGAGGCGGACATAATGGGTTAGGG 840
DB 781 TTAACCAACGATGACGATGAGTGGGCTTGAACGAGGCGGACATAATGGGTTAGGG 840
QY 841 AGCGACCCACGAGATGATTAAGGCGCCACTACACTGGCCGTTTCCTCATCAGTGGCC 900
DB 841 AGCGACCCACGAGATGATTAAGGCGCCACTACACTGGCCGTTTCCTCATCAGTGGCC 900
QY 901 CAAACCCCAATGTGAACGATGACCGGCTGATGCGGCGCGGTAAACATGGAATCTTTCG 960
DB 901 CAAACCCCAATGTGAACGATGACCGGCTGATGCGGCGCGGTAAACATGGAATCTTTCG 960
QY 961 AATCTAGCTTTGAGGTGCTCCGGAATTTGAACATCTTCCACGCTTAAGGACATGACGATA 1020
DB 961 AATCTAGCTTTGAGGTGCTCCGGAATTTGAACATCTTCCACGCTTAAGGACATGACGATA 1020
QY 1021 TCTACAGCAAGTACTAGTATCATCATGATGATTAACCAACGACACTGGAGAAACGTGTGG 1080
DB 1021 TCTACAGCAAGTACTAGTATCATCATGATGATTAACCAACGACACTGGAGAAACGTGTGG 1080
QY 1081 ATGCTCTCAAGAAATCAGGCACTTGTCTCATCTCAGCTATCAGCTCAGCGCAGTTTG 1140
DB 1081 ATGCTCTCAAGAAATCAGGCACTTGTCTCATCTCAGCTATCAGCTCAGCGCAGTTTG 1140
QY 1141 TCGCTGTACAGCTTAAGGAATTTGTTAAGCTTCTGTGACATCTCTCAAGAGAACTAC 1200
DB 1141 TCGCTGTACAGCTTAAGGAATTTGTTAAGCTTCTGTGACATCTCTCAAGAGAACTAC 1200
QY 1201 GATCAGACGTGATCCGAGAGCGGATGATCAGCATGCGCTATCATGCTTAAGACGCTGAAGA 1260
DB 1201 GATCAGACGTGATCCGAGAGCGGATGATCAGCATGCGCTATCATGCTTAAGACGCTGAAGA 1260
QY 1261 ATAACTAGATGCTCTGCTGAGACATTTTGAAGACACTGATTAATTAATACGAACA 1320
DB 1261 ATAACTAGATGCTCTGCTGAGACATTTTGAAGACACTGATTAATTAATACGAACA 1320
QY 1321 GCGCGAAGGTCAATGATCCGCTTCAACATAGCTCTGAAGTATATCTTAAGTATAC 1380
DB 1321 GCGCGAAGGTCAATGATCCGCTTCAACATAGCTCTGAAGTATATCTTAAGTATAC 1380
QY 1381 ATGACCAAGACCTGCTCAAGTCTTCAACAGACACTGTAATCAATCAAGTCAAGGACA 1440
DB 1381 ATGACCAAGACCTGCTCAAGTCTTCAACAGACACTGTAATCAATCAAGTCAAGGACA 1440
QY 1441 TAAGTCCACACTGTGTGAGCTGATGAGTGTCTCTTCAAGAGATGGCAGAGGCGT 1500
DB 1441 TAAGTCCACACTGTGTGAGCTGATGAGTGTCTCTTCAAGAGATGGCAGAGGCGT 1500
QY 1501 TGAAGGAATGCAACCGTACTAAGGACACCTTTAAACAAATCCATTGGGATGACACT 1560
DB 1501 TGAAGGAATGCAACCGTACTAAGGACACCTTTAAACAAATCCATTGGGATGACACT 1560
QY 1561 GCGATGACAGCGCGGCACTTCCAGATACGCTTATGGGCTTTCAAGGCGTCACTTCCAGAGC 1620
DB 1561 GCGATGACAGCGCGGCACTTCCAGATACGCTTATGGGCTTTCAAGGCGTCACTTCCAGAGC 1620
QY 1621 TGGCGGATCAAAATATATGAAACATTAGACATAGCTGCCACGCGGCACTTGAAGAGGAAAC 1680
DB 1621 TGGCGGATCAAAATATATGAAACATTAGACATAGCTGCCACGCGGCACTTGAAGAGGAAAC 1680
QY 1681 GAGAGGGCGGCGGAGAGAGAACTGTACTGGGACTGGGACTGACCTTGAACGAGAC 1740
DB 1681 GAGAGGGCGGCGGAGAGAGAACTGTACTGGGACTGGGACTGACCTTGAACGAGAC 1740
QY 1741 GCACTGTATCCCGATTTGGCCGAACACTGGAACCTTGAACCCCAAGCCTTGTATGA 1800
DB 1741 GCACTGTATCCCGATTTGGCCGAACACTGGAACCTTGAACCCCAAGCCTTGTATGA 1800
QY 1801 GATCCATTTACGCGGTGAGACACTGCGGCTGCTCAACGAGCCAAAGTTAAGGCGCAATATA 1860
DB 1801 GATCCATTTACGCGGTGAGACACTGCGGCTGCTCAACGAGCCAAAGTTAAGGCGCAATATA 1860
QY 1861 CACTATATTTCCAGGCAAGAAACCTTTAGGAGCTAATAATTTCCACAGGACATGATGA 1920

```

Db 1861 CACTATTTCAGGAGAAAGAAACCTTTAGACCTAAATTAATTCACACGAGATGATGA 1920
 Qy 1921 CAGGACGACGACGAGTGGATCACTGACCCAGACCTCGCCTGAATTCCTAATAGCGGTGCA 1980
 Db 1921 CAGGACGACGACGAGTGGATCACTGACCCAGACCTCGCCTGAATTCCTAATAGCGGTGCA 1980
 Qy 1981 CACGACTACAAACGCGGGATCGTTTACCAACGCGCCCGGGAGACGGGACAGATATTC 2040
 Db 1981 CACGACTACAAACGCGGGATCGTTTACCAACGCGCCCGGGAGACGGGACAGATATTC 2040
 Qy 2041 AGTCACGACGACGAGTGGAGATCCACTCGCCACAGACAAATAGAGATCACTAGCGTG 2100
 Db 2041 AGTCACGACGACGAGTGGAGATCCACTCGCCACAGACAAATAGAGATCACTAGCGTG 2100
 Qy 2101 GTATTGGTAATTAATCTATCGTGGGGCCACTGGCGCCATCCAAAAGGCTCCGGAATAC 2160
 Db 2101 GTATTGGTAATTAATCTATCGTGGGGCCACTGGCGCCATCCAAAAGGCTCCGGAATAC 2160
 Qy 2161 CCCGAGACACGACGCTCTAGGGAAACGAGTCCAATAGTCAAGTGGGTGCTTATGA 2220
 Db 2161 CCCGAGACACGACGCTCTAGGGAAACGAGTCCAATAGTCAAGTGGGTGCTTATGA 2220
 Qy 2221 AACGACGATGTACTCTACAGGTGGGGGTCTCGACGTACGCCGAGAGAAACACCCAG 2280
 Db 2221 AACGACGATGTACTCTACAGGTGGGGGTCTCGACGTACGCCGAGAGAAACACCCAG 2280
 Qy 2281 TAAAGCAATCGCGCGCGGACGACGTCTGGGGCAATCCGTGAGACAGAACTATACATTAG 2340
 Db 2281 TAAAGCAATCGCGCGCGGACGACGTCTGGGGCAATCCGTGAGACAGAACTATACATTAG 2340
 Qy 2341 GCGTCGAGATGATGAGACAGCCAGACTATGTTCCGGGGACTACATGCGACGCGGAA 2400
 Db 2341 GCGTCGAGATGATGAGACAGCCAGACTATGTTCCGGGGACTACATGCGACGCGGAA 2400
 Qy 2401 TCGCGATGGGTGAGAAAGCTCATGGACGTGATGAGTGTGATGACATCGACTCCGAGGCCA 2460
 Db 2401 TCGCGATGGGTGAGAAAGCTCATGGACGTGATGAGTGTGATGACATCGACTCCGAGGCCA 2460
 Qy 2461 GTTCGTGTGTTCAGAAAGCTCTTTTCAGTCCAGCTACACTAGAGGTAAATGAATGAACT 2520
 Db 2461 GTTCGTGTGTTCAGAAAGCTCTTTTCAGTCCAGCTACACTAGAGGTAAATGAATGAACT 2520
 Qy 2521 ACTCACTTAGCGGAGCCACACCCGCTTGAATGAGCAACGACGCGGCGCATTTGAGC 2580
 Db 2521 ACTCACTTAGCGGAGCCACACCCGCTTGAATGAGCAACGACGCGGCGCATTTGAGC 2580
 Qy 2581 ACATGAGACGATTAATTCAGTTCTGCGCATCGACGATTTGTTGAAAGAAAGATGGCC 2640
 Db 2581 ACATGAGACGATTAATTCAGTTCTGCGCATCGACGATTTGTTGAAAGAAAGATGGCC 2640
 Qy 2641 TGAATCAGCTTCAACAGTATCTGGCGGATGGAAGAGCTCACTCGACGAACTTAAT 2700
 Db 2641 TGAATCAGCTTCAACAGTATCTGGCGGATGGAAGAGCTCACTCGACGAACTTAAT 2700
 Qy 2701 GGGTCTGACATGTTCCGACAGATGTTCAATGATACACACCAAGTGTACTCACTGT 2760
 Db 2701 GGGTCTGACATGTTCCGACAGATGTTCAATGATACACACCAAGTGTACTCACTGT 2760
 Qy 2761 TTTTGGACACGGTCAACGAACTGATTTCTGGTTCATGCGAAACGAACTTCAAGAAATGGCT 2820
 Db 2761 TTTTGGACACGGTCAACGAACTGATTTCTGGTTCATGCGAAACGAACTTCAAGAAATGGCT 2820
 Qy 2821 CTTCATCTTTTGAAGCGGTTGTTCAACAAATTGGGACCGATCTACCAATTCGATGC 2880
 Db 2821 CTTCATCTTTTGAAGCGGTTGTTCAACAAATTGGGACCGATCTACCAATTCGATGC 2880
 Qy 2881 ACAGCAAGATTTGGAAGACGCTACAGTGGTCCACGAATATTTCCCAACGACAGTTGCAAC 2940
 Db 2881 ACAGCAAGATTTGGAAGACGCTACAGTGGTCCACGAATATTTCCCAACGACAGTTGCAAC 2940
 Qy 2941 TTAAGAACTAATTTAGATCAATTCGATTTCTAATAACGCACTACCAAGACGGCA 3000
 Db 2941 TTAAGAACTAATTTAGATCAATTCGATTTCTAATAACGCACTACCAAGACGGCA 3000

Db 2941 TTAAGAACTAATTTAGATCAATTCGATTTCTAATAACGCACTACCAAGACGGCA 3000
 Qy 3001 TCGCAATTCGCGCTTCTTAACGATCTGGCCAAATATATCTGTAAAGACGCACTTTC 3060
 Db 3001 TCGCAATTCGCGCTTCTTAACGATCTGGCCAAATATATCTGTAAAGACGCACTTTC 3060
 Qy 3061 CCACGACACGAGCCGAGCGGAGCGGACCGGCTTAAGCTGGCCAGCTGGCGGCG 3120
 Db 3061 CCACGACACGAGCCGAGCGGAGCGGACCGGCTTAAGCTGGCCAGCTGGCGGCG 3120
 Qy 3121 ATCAGAAGTCGATGAGCTGGCTCCAGGACGAGAGCTGGCTTATGAGCCCTGTATTAAC 3180
 Db 3121 ATCAGAAGTCGATGAGCTGGCTCCAGGACGAGAGCTGGCTTATGAGCCCTGTATTAAC 3180
 Qy 3181 TGAATACCCCGCAATGACCTTTTATCGGCCACCTGGCCAAAGTATATAGAACTCTG 3240
 Db 3181 TGAATACCCCGCAATGACCTTTTATCGGCCACCTGGCCAAAGTATATAGAACTCTG 3240
 Qy 3241 CCCGATCTGTGATCATTTCCGACATGAGGCGGCAAGCAAGTTGCATTCGGGTGCCA 3300
 Db 3241 CCCGATCTGTGATCATTTCCGACATGAGGCGGCAAGCAAGTTGCATTCGGGTGCCA 3300
 Qy 3301 ATTGCGCTAGTACTCTCATTTGAGACGATGACAGTCCCAAGCTTTGCAAGTCCCTCTG 3360
 Db 3301 ATTGCGCTAGTACTCTCATTTGAGACGATGACAGTCCCAAGCTTTGCAAGTCCCTCTG 3360
 Qy 3361 TGGGTCAATTTGCTCGCTTCAGAGCCACACACCAACTTAAGATCAGCTCTACTATC 3420
 Db 3361 TGGGTCAATTTGCTCGCTTCAGAGCCACACACCAACTTAAGATCAGCTCTACTATC 3420
 Qy 3421 CAGCCTCCCGCACTCTCCGTGAGACGAGAGCTCTTTTCCGAGCTGACATTC 3480
 Db 3421 CAGCCTCCCGCACTCTCCGTGAGACGAGAGCTCTTTTCCGAGCTGACATTC 3480
 Qy 3481 CAGCCTCCCGCACTCTCCGTGAGACGAGAGCTCTTTTCCGAGCTGACATTC 3480
 Db 3481 CAGCCTCCCGCACTCTCCGTGAGACGAGAGCTCTTTTCCGAGCTGACATTC 3480
 Qy 3481 AGCAACACATTCAGAAAGCTCGAGAGAGATCCGCACTGCTTGGCGGTGATACCA 3540
 Db 3481 AGCAACACATTCAGAAAGCTCGAGAGAGATCCGCACTGCTTGGCGGTGATACCA 3540
 Qy 3541 CGGCGCTGGCCCAATGGCTTCAATGGACCTTGCACTGATACGATGAGGCGCAACAG 3600
 Db 3541 CGGCGCTGGCCCAATGGCTTCAATGGACCTTGCACTGATACGATGAGGCGCAACAG 3600
 Qy 3601 ATTGCTGCAATCCCTGCTTCAACTCAACGCAAGCAATCGCGGCAACTACCCAGT 3660
 Db 3601 ATTGCTGCAATCCCTGCTTCAACTCAACGCAAGCAATCGCGGCAACTACCCAGT 3660
 Qy 3661 CAATATCACCTGAGTCAGCAACAAATGAGCTGATTAATCTGAGCGGGAAGGACCACTC 3720
 Db 3661 CAATATCACCTGAGTCAGCAACAAATGAGCTGATTAATCTGAGCGGGAAGGACCACTC 3720
 Qy 3721 AGAAGCGCAAGTCACCAACTGACGATGCAAGTGTGATCAAGTCTCGATTAATATGCTG 3780
 Db 3721 AGAAGCGCAAGTCACCAACTGACGATGCAAGTGTGATCAAGTCTCGATTAATATGCTG 3780
 Qy 3781 AAAATGAGACCTGATATGAGGCGGACCAACCTGATGGAAGGCAAGTGTGCTGCT 3840
 Db 3781 AAAATGAGACCTGATATGAGGCGGACCAACCTGATGGAAGGCAAGTGTGCTGCT 3840
 Qy 3841 TGAAGCTAACAAAGATCAGCCCGTCAAGTGTCTTCAAGCTCACTTAACCTGGGA 3900
 Db 3841 TGAAGCTAACAAAGATCAGCCCGTCAAGTGTCTTCAAGCTCACTTAACCTGGGA 3900
 Qy 3901 TTTGATCAAGGCGGAAACTGTGAGCTGCCCAATAGCACTTTAGATGATATGCGAA 3960
 Db 3901 TTTGATCAAGGCGGAAACTGTGAGCTGCCCAATAGCACTTTAGATGATATGCGAA 3960
 Qy 3961 TGTCTCTTAACATTTGAGAGCGGACGATAGGAGGATGATGCTGCTGGCCGTCGACGTC 4020
 Db 3961 TGTCTCTTAACATTTGAGAGCGGACGATAGGAGGATGATGCTGCTGGCCGTCGACGTC 4020
 Qy 4021 TCAAGTAAGTATATGAGAGCAACAAATGCTGCAACATGAGATGCACTTTAGAGCTGA 4080
 Db 4021 TCAAGTAAGTATATGAGAGCAACAAATGCTGCAACATGAGATGCACTTTAGAGCTGA 4080

QY 4081 TTTTGTGAAGATCATCCAGTGTCTATCAACACAGCAGAGAGGCTTTGGGGATATCGACT 4140
 DB 4081 TTTTGTGAAGATCATCCAGTGTCTATCAACACAGCAGAGAGGCTTTGGGGATATCGACT 4140
 QY 4141 CGATGATACCAAGATACACACATCTCTGGATCTGTCCATTCACATTCGCAATC 4200
 DB 4141 CGATGATACCAAGATACACACATCTCTGGATCTGTCCATTCACATTCGCAATC 4200
 QY 4201 CGGTAAATTCGACGCGGTGAGTTTCCACAAAATTTGTGGCCATCAAGATCCGTGAGAG 4260
 DB 4201 CGGTAAATTCGACGCGGTGAGTTTCCACAAAATTTGTGGCCATCAAGATCCGTGAGAG 4260
 QY 4261 TGACCGAACACACATGTTTGGAGATCACAGACGCCCATTTGGATTTCCCAATC 4320
 DB 4261 TGACCGAACACACATGTTTGGAGATCACAGACGCCCATTTGGATTTCCCAATC 4320
 QY 4321 TGGGGCCATCGGCGGACGACACGCAATGATGTTCCGAAAGCTTCCTTCGCAATC 4380
 DB 4321 TGGGGCCATCGGCGGACGACACGCAATGATGTTCCGAAAGCTTCCTTCGCAATC 4380
 QY 4381 TCAAGCTGTACTTTGTTCTGGGCGAGAGAGTCAAGCCGAAAGCTGTCACTGCTAAATC 4440
 DB 4381 TCAAGCTGTACTTTGTTCTGGGCGAGAGAGTCAAGCCGAAAGCTGTCACTGCTAAATC 4440
 QY 4441 CCAGCAAGTTAGGCTCTCTCAACGTTACATCGAAGACGCGAACTGCATCAGTGGGG 4500
 DB 4441 CCAGCAAGTTAGGCTCTCTCAACGTTACATCGAAGACGCGAACTGCATCAGTGGGG 4500
 QY 4501 GAGGAAGCTTACAAAGAACTCTCCGGGATGTCGTCTGATGTTGGGGAGCCCTTAA 4560
 DB 4501 GAGGAAGCTTACAAAGAACTCTCCGGGATGTCGTCTGATGTTGGGGAGCCCTTAA 4560
 QY 4561 TAGGATTCCTGCTGTCGCCACCAACAACAAGACAGAGCGGTTGCTTCCCTGGGCTG 4620
 DB 4561 TAGGATTCCTGCTGTCGCCACCAACAACAAGACAGAGCGGTTGCTTCCCTGGGCTG 4620
 QY 4621 AGAAGAAAGTGAAGAGCGCGGACATTAATATATATATATATATATATATATATATAT 4680
 DB 4621 AGAAGAAAGTGAAGAGCGCGGACATTAATATATATATATATATATATATATATATAT 4680
 QY 4681 TAACTAAT 4740
 DB 4681 TAACTAAT 4740
 QY 4741 ATGCAAAATCGCGCCGGAAGAAAGTCAAGGCGCATTCGCCATTAATGTCAGAGAAAA 4800
 DB 4741 ATGCAAAATCGCGCCGGAAGAAAGTCAAGGCGCATTCGCCATTAATGTCAGAGAAAA 4800
 QY 4801 CTATTTATACATAACCGGAATTAAGCGAGTAATTCGCAAGAGTACTTAAATTTGAG 4860
 DB 4801 CTATTTATACATAACCGGAATTAAGCGAGTAATTCGCAAGAGTACTTAAATTTGAG 4860
 QY 4861 CCAACTCGGCAATCTCTCTCTCAATCAGTGTGAGCTTTTTCAGAGCTCTTAATACTA 4920
 DB 4861 CCAACTCGGCAATCTCTCTCTCAATCAGTGTGAGCTTTTTCAGAGCTCTTAATACTA 4920
 QY 4921 TTAACGCGATTCCTGTTTATTTGTAAGCTTATACCGCTTTTATGGAACACAGAG 4980
 DB 4921 TTAACGCGATTCCTGTTTATTTGTAAGCTTATACCGCTTTTATGGAACACAGAG 4980
 QY 4981 TGTGCTTTTCTACCTTTGTTTCAAGCGCATTTGTAATCTTATTTGTTCCGCTCAACTA 5040
 DB 4981 TGTGCTTTTCTACCTTTGTTTCAAGCGCATTTGTAATCTTATTTGTTCCGCTCAACTA 5040
 QY 5041 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5100
 DB 5041 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5100
 QY 5101 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145
 DB 5101 AATTTCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145

RESULT 2
 ADB62173
 ID ADB62173 standard; cDNA; 3153 BP.
 XX
 AC ADB62173;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human cDNA encoding clone CTONG20032930.
 KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
 KW tissue regeneration; cell regeneration; membrane protein;
 KW signal transduction-related protein; transcription-related protein;
 KW osteoporosis; neurological disease; cancer; tumour.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 352..3102
 FT /*lag= a
 FT /product= "Clone CTONG20032930 protein"
 XX
 PN EPI308459-A2.
 XX
 PD 07-MAY-2003.
 XX
 PF 28-MAR-2002; 2002EP-00007401.
 XX
 PR 05-NOV-2001; 2001JP-00379298.
 PR 25-JAN-2002; 2002US-00350978.
 XX
 PA (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 PI Iisogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamehika I,
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Nasuno Y;
 XX
 DR WPI; 2003-450961/43.
 DR P-PSDB; ADB64143.
 XX
 PT New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.
 PS
 PS Claim 1; Page; 222pp; English.
 XX
 CC The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesizing the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC genes may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX
 SQ Sequence 3153 BP; 883 A; 701 C; 767 G; 802 T; 0 U; 0 Other;

Query Match 2.9%; Score 149.8; DB 10; Length 3153;
 Best Local Similarity 47.1%; Pred. No. 8.3e-32;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

232 CATGCTGACGGGACCACTTTAAGATTGACAAAAGTCCCTGGAGCGTTCTGGAGC 291
 521 CTGGGATGAATCTAGCAATTAACAAGGTGTTCTGCTGGGACATGACATCCGTCCGCC 580
 232 TAAATAGGATTTGGGACCGATTTTAATGACATACACGGCTACCGTTCCGACATNGA 351
 581 TGGTACCCGGCTGACAGATCGGTTCAAGCCGACAGTGGCACAAGTGTGCCAAGCTPA 640
 352 TCGATCGCTGGGAGACACAGGAGACACAGTCCGCGAGAGGCGCAACTTCTGCTCGCG 411
 641 TTAGACAGATAGAGATGCTTAAGACTCTGTAGAGGAGACAGAACCAATCTGCTGCTPA 700
 412 ACCTGATGAGACACAGATGCTTCCGCCAGCGCTGATGACAAAGCTGGCCACTAGCT 471
 701 AGATCATGATCAAGCTGCTAATCCCACTAGCTATGAGACAGAAATGCT-----TGGAG 754
 472 GCTTGAAGACAGAAAGCCAGAGTCCGCGAGAGAGTTCCTGACAGATTTGAAAGCTC 531
 755 GCTTCAAAACAAAGAAATTTCCGCTACTCGAAGAGCATCTGCTCTGCTTATAGCAAC 814
 532 TTCTAGATGACGACCCAGCACTTATGTTGCGCTCTATATACCAACAGTTTGTGAC 591
 815 TCAATGCTCTGAGACACAGACTTTTAACACTAAGCAAGATTTGTGCAACATATATGCA 874
 592 TTCTGAGAGATCCACAGTTAATGTAGAGGAGCGGCGCATCCAAACGCTAGTGAATCT 651
 875 TACTTGGAGATCAAAACAGCGTTCGAGATGACAGCAATAAACAGCTTAGTGAATTT 934
 652 AAGACATGATGAGGATGATGCTCCGCCAGACCTCCGTGCGATGACAGATTTCCGCT 711
 935 AAGACATGATGAGGATGATGCTCCGCCAGACCTCCGTGCGATGACAGATTTCCGCT 991
 712 CGAAATTTGCTATGTTGAGCAAAAAGTTGACCAAGGCTCAACAGAGAGGCTTACTGCTAC 771
 992 CCGGTTGATGATTTTATTTTACAAAATTTGATGAGAGTCCAGAAATGTGAAACATGATAC 1051
 772 CTTCAGCCCTTAAAAACAGAAATGGCAATGAGTGGCTTTGACAGAGGCC-GACAAATAT 830
 1052 AATTCGCAATGATTAATAAATTTGACAGATGAAGATTTCTGTGATGTTAACAACCTTCT 1111
 831 GGTGAGGAGGACGACCCACAGAGATGATTAAGCGGCCACTACCTCGGCCCTTTCGTCA 890
 1112 CTGCTAGTTCTATCATATATCCAAAGGCTCCACCAAGTTCTCGAGAAACGTTGGAATGGAA 1171
 891 -----TCACTGCGCCCAAAACCAATGTGAACGATGTGACCGGTGATG 933
 1172 CCAACCGCGCGGCTGTTATTCACCTTGTGATTCAGTCTTCAGTCTGCAAAAAGAGAG 1231
 934 CCGGCGCGCTAAACCATGATCTTTTGAATCTAGCTTTGAGTGTCCGCAATTGACA 993
 1232 CTGGTGTCTTTGATGAAGAGATTTTATTAAGCAATTTGATGATGATCTGATGATCAGA 1291
 994 TCTTCACGCTTAAGACATGAGAGATATCTACAAGCAAGTACTAGTATCATCAAGTAA 1053
 1292 TTTATTCACGCGGAGCTTTGAGAGATTCATAAAACAAATTTAGGAAATTTATCTGATG 1351
 1054 AAAAGCAGACTGGGAGAAACGTGTGATGCTCTCAAGAAAGATCAGGCAATGCTCATTC 1113
 1352 ACAAGCATGTTGGAGCAGAGAGTAAATGCTTAAAGAAAGATTAATCTTTACTTT--- 1409
 1114 TCAGTATCACTCAGCGCAGCTTTGTGCTGTGATGATGCTTAAAGAAATGCTTAAAGCT 1173
 1410 ---GGCTGTGCTGTGATGATGATGATCTTTTCAACATTTGCTCTTTTGGATGAG 1465
 1174 TGTGACATCTCAGAGAGAACTACATCAAGGTGATCCGCGAGGCTGTGATCACA 1233
 1466 CTTTAAATCTCTGTGATGAGACTCGGCTCAGGATGATGCGGAGGCTTGTATCAGCT 1525

1234 TCGCTACATGCTTAAGACGCTGAGAAATAAATCTAGATGCTTCTGCTGAGCAATTTGG 1293
 1526 TGGGACATCTGCTCATCAGTTCTGGGAAATAGTTGACCATGAGCTGAGCCATTTATGC 1585
 1294 AGACCTGATTAATTTAATAACAGAACGCGGAGTCAATTCAGCTTCCACAAATAG 1353
 1586 CAATCTTTTATTTAATTTTCAAAACAGTCCCAAAATTTATGCCACATCTGCTGTAG 1645
 1354 CTCTGAGTATATGATTAATGATTAACATGACCAAAAGCTCTCAAGATCTTACACAGCA 1413
 1646 CTGTTAGTTATTTATTTGAGACACACATCCCTAGTTAATCTGTCTATTAACAAGCA 1705
 1414 CTCTGATCAATCAAAAGTTCAGAGACATTAAGTCCACCTGTGAGCTGATGAGCTGC 1473
 1706 ---ACTGTACCTCTAAGCTGTGCTGCAAGTAAAGCGCTGTTTGAATTTTGAATTTGC 1762
 1474 TCTTCGAGAGTGGGACAGAAAGCGTTGAAAGAAATGCCACGCTACTAAGAGACACT 1533
 1763 TTTTACAGAAATGACAGACATTTCACTAGAACACACATATCAATTTAGCTGAAACAA 1822
 1534 TAAATAATTCATTTGGCGATGACAGCTGATGACAGCGCCCATTCAGATACGCTTAT 1593
 1823 TAAAGAGGAATACATGATGCTGATTCGAAAGCAAGATAGAAAGCCGAAATGTTACT 1882
 1594 GGGCTTCAAGCGCTCACTTTCCAGAGCTGGGGATCAATAATATGGAACATTTAGCATAG 1653
 1883 GGGGTTTCCACAGTCACTTCCAGAGAGAGAGACACTTGTACCAACCTTGGAGTCT 1942
 1654 CTGCCAGCGCGC 1666
 1943 CTTACAGAAAGC 1955

RESULT 3
 ADR83507
 ID ADR83507 standard; DNA; 7888 BP.
 AC ADR83507;
 XX 02-DEC-2004 (first entry)
 XX
 DE Human multiple asters 1 DNA, target gene of miRNA.
 XX
 KW human; ds; miRNA; microRNA; ontogenesis; cell therapy; cancer;
 KW immune disease; nerve disorder; amyotrophic lateral sclerosis;
 KW Parkinson's disease; Alzheimer's disease; inflammatory disease;
 KW siRNA silencing precursor; cytosolic; immunosuppressive; nootropic;
 KW neuroprotective; antiinflammatory; immunotherapy; multiple asters 1.
 KW
 OS Homo sapiens.
 XX
 PN W0204076622-A2.
 XX
 PD 10-SEP-2004.
 XX
 PF 10-FEB-2004; 2004WO-JP001433.
 XX
 PR 10-FEB-2003; 2003US-0445829P.
 XX
 PA (NABD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 XX
 PI Taira K, Kawasaki H;
 XX
 DR WPI; 2004-653393/63.
 XX
 PT Modulating expression of a target gene in a cell, for treating cancer, an
 PT immune disease, or a nerve disorder, comprises introducing into the cell
 PT a polynucleotide that forms a duplex region with an mRNA transcribed from
 PT the target gene.
 XX
 PS Claim 9; SEQ ID NO 409; 865bp; English.
 XX
 CC This invention relates to a novel method for modulating the expression of

CC a target gene in a cell. Specifically, it refers to the introduction into
 CC a cell of a polynucleotide that forms a duplex region with an mRNA
 CC transcribed from the target gene, where the duplex region comprises a
 CC mammalian miRNA target region i.e. a non-coding microRNA (miRNA) that
 CC regulates mRNA at a post-transcriptional level. The present invention
 CC describes a method for controlling the onset of a mammal, function of a
 CC mammalian cell, differentiation of a mammalian cell or viability of a
 CC mammalian cell in the post-transcriptional phase, which comprises
 CC introducing a plasmid vector comprising a promoter and nucleic acid
 CC molecule expressing an miRNA or siRNA silencing precursor to the miRNA.
 CC Accordingly, it provides a cell therapy method for treating cancer,
 CC immune disease, nerve disorder (e.g. Amyotrophic lateral sclerosis,
 CC Parkinson's disease, or Alzheimer's disease) or an inflammatory disease
 CC by introducing into the cell the miRNA, siRNA silencing precursor to the
 CC miRNA or the plasmid vector. As such, they can be developed into
 CC pharmaceutical compositions that exhibit cytostatic, immunosuppressive,
 CC neurotrophic, neuroprotective and anti-inflammatory activities and hence can
 CC be used for immunotherapy. This polynucleotide sequence is a human target
 CC gene whose expression is modulated by miRNAs of the invention.

XX Sequence 7888 BP; 2129 A; 1876 C; 1858 G; 2025 T; 0 U; 0 Other;

Query Match 2.9%; Score 149.8; DB 13; Length 7888;
 Best Local Similarity 47.1%; Pred. No. 1.5e-31;
 Matches 685; Conservative 0; Mismatches 732; Indels 36; Gaps 6;

QY 232 CATGCTGACGGGACGACCTTTAAGATTGCAAAAATCCCTGGAGCGTTCTCGAGC 291
 DB 335 CTGGGTGTAAGTCTAAGCAATTAAGAGTGTTCTGCTGGGATGACATCTGTCCGCC 394
 QY 292 TAAATAAACGATGGGACGCAATTTAATGCAATACCGCTACCGTTTCCCAATGTA 351
 DB 395 TGGTGACCCGGCTGAGGATCGTTCAAGGGCGAGATCGGACAGTCTGCAAGTCTAA 454
 QY 352 TCGATCGCTGGGAGACAGGAGACAGTCCGGAGAGGCGCAATTTGCTGGCGCG 411
 DB 455 TAGACAGCTGAGAGATCTTAAAGCTTGTGAGGAGACAGACCAATCTGTCTTAA 514
 QY 412 ACCATATGAGACAGAGATGCTTCCGCCCAAGCGCTGATGACAAAGCTGCGCAATGCT 471
 DB 515 AGATCATGATGATCAAGCTCTAATCCCAAGTATGAGACAGATGCT-----TGGAG 568
 QY 472 GCTTCAAGCAAGAAGCCCAAGGTGCGGAGGACTTCTTCAAGCATTTGTGAAGCTC 531
 DB 569 GCTTCAAGCAAGAAGTTCGTAATCTCGAAGAGGATCTGTCTTCAATTAACCAAC 628
 QY 532 TTCAATGATGAGGACCCAGAGCTTATGTTGCGTATATACCAATTTGTGAC 591
 DB 629 TCAATGCTTGGAGACAGACTTTAACAATAAGATTTGTGCAATATATGCAACT 688
 QY 592 TTCTCGAGATCCCAAGTATATGAGGAGGCGGCATCAACCGTATGGAATCT 651
 DB 689 TACTTGGAGATCCAAAGCCAGGTTGAGATGACGAATTAACGCTTATGGAATTT 748
 QY 652 ACAAGCATGAGGGGATGATGGCCCAAGCTTCCTGTCGATGAGCATGTTCTGCT 711
 DB 749 ACAAGCATGAGGAGATGATGGGCAAGTCTCAGTAA---AAGATTGCAACAGT 805
 QY 712 CGAATTTGCGATATGTTGAGCAAAAGTTCGACCAAGTCAACAGAGGCTTACTGCTAC 771
 DB 806 CCGGTTTAAATGTAATTTTAAACAAATTTGATGAAGTCAAGAAATCTGGAACATGATAC 865
 QY 772 CTTCAGCCCTTAAACACGAATGGCAATGAGTGGCTTGGAGAGGCC-GACAAATTT 830
 DB 866 AATCTGCAAAATGATTAATAATTTGACGATGAAGATTTGTGATGATTAACAGACTTCT 925
 QY 831 GGGTTGAGGAGGACCAACCAAGATGATTAAGCGGCACATCACTCGGCGCTTTCTCA 890
 DB 926 CTGCTAGTTTACATCATCAAGGCTCCACCAAGTTCTCGAGAAAGAACTTGAATGGAA 985
 QY 891 -----TCACTGGCGCCAAACCAATGTAAGATGATGACCGGTATG 933
 DB 986 CCACCCCGCGGCTTGTTCATCATCACTTGTGATTCAGTCTTCAAGTCAAAAGAGAG 1045

QY 934 CCGGCGCGGTAAACCATGGAATCTTTGCAATCTAGCTTTGAGTGTGTCGCCAATTTGACA 993
 DB 1046 CTGGGTGTTGATGAAGAGATTTTATTAACATTTGATGATGTACCTGTAGTACAGA 1105
 QY 994 TCTTCCACGTTAAGACATGAGATGATATATCAAGCAAGTATGATCATCGATATA 1053
 DB 1106 TTTATTTCCAGCCAGACCTTTGAGGAATCTCAATTAACAAATTTAGGAAATTTATCTATG 1165
 QY 1054 AAAACGAGACTGGGAGAAAAGTGTGATGCTCTCAAGAAATCAGGGCATTTGCTCATTC 1113
 DB 1166 ACAAGCATGATTTGGAGACAGAGTAAATCTTAAAAAGATTAAGATCTTTACTTT-- 1223
 QY 1114 TCAAGTATCACTCAAGCCGACGTTGTGCTGTGATCACTTAAAGAAATTTGCTTAACT 1173
 DB 1224 ----GGTGTGCTGTGATGATGATTAATCTTTCAACATTTGCGCTTTTGGATGAG 1279
 QY 1174 TCGTGACATCTTCAGAGAGAACTACATCAAGGATTCGGAGGCGGTGATCAACA 1233
 DB 1280 CTTTAAACTCTCTGCTAAGGACCTGCGGTCTCAAGGATGCGGAGCTGTATACGT 1339
 QY 1234 TCGCTATCATCTTCAAGACCGTGAAGAAATTAATAGATGCTTCTGCTGAGCATTTTGG 1293
 DB 1340 TGGGCACTCTGTATCATGATCTTGGGAAATAGTTGACATGAGCTGAAGCATTTATGC 1399
 QY 1294 AGCACTGATTAATTTAATACAGAACGCGGAAGCTCATTCGATCCGTTTCAATAG 1353
 DB 1400 CAATATCTTTAATTTAATTTCAAAACAGTCCAAAATTTATGGCCACATCTGTGTGTAG 1459
 QY 1354 CTGTAAGTATATCATTAAGTATACATGACCAAAAGCTGCTCAAGATCTACACAGCA 1413
 DB 1460 CTGTAGGTAATTAATTTGCGACACACATCTTCTAGTTAATCTGTATTAACAGCA 1519
 QY 1414 CTCTGATCAATCAAAAGTTCAGAACATTAAGTTCACATCTGTGAGCTGATGCTGTC 1473
 DB 1520 ---AGTACTCTTAAGTCTGTGAGCTGTGAGAGGCGCTTTTGAATTTTATGATTTGC 1576
 QY 1474 TCTTCAAGAGTGGACAGACGAGCGTTGGAAGAAATGCCACCTTCTAAGGACACT 1533
 DB 1577 TTTTACAAAGATGACAGACATTTCACTAGAACAGACATATGATTTAGCTGAACAA 1636
 QY 1534 TAAAAAATCCATTTGGAGTACAGACTGCGATGACGCGCCATTCAGATACGCTATT 1593
 DB 1637 TAAAGAGGAAATACATGATGATTTCCAGACAAAGATTAAGACCAAGAAATGTTACT 1656
 QY 1594 GGGCTTTCAGGCTCACTTTCCAGAGCTGCGGATCAATATATGGAACAATTAGACATAG 1653
 DB 1697 GGGTTTCCACAGTCACTTTCAGCAGAGAAACAGACATTTGACACACCTTGGAGTCT 1756
 QY 1654 CTGCCACGCGCGC 1666
 DB 1757 CTTACCGAAGAGC 1769

RESULT 4
 AAC64201
 ID AAC64201 standard; cDNA; 6245 BP.
 XX
 XX AAC64201;
 DT 21-FEB-2001 (first entry)
 XX
 XX Human polliosis-associated gene 373 cDNA, SEQ ID NO:1.
 DE Human polliosis-associated gene 373; IgE; Immunoglobulin E;
 KW cedar pollen allergy; T-cell; reduced expression; detection; diagnosis;
 KW drug screening; allergic disease; ss.
 XX Homo sapiens.
 XX OS
 XX NC020005046-A1.
 XX FN
 XX PD 02-NOV-2000.

XX 26-APR-2000; 2000WO-JP002730.
 XX 27-APR-1999; 994P-00120489.
 PR (GENO-) GENOX RES INC.
 XX Nagasu T, Sugita Y, Kashiwabara T, Oshida T, Obayashi M, Gunji S;
 PI Obayashi I, Imai Y, Yoshida N, Ogawa K, Matsui K;
 XX MPI; 2000-687339/67.
 DR P-PSDB; AAB29634.
 PT Pollinosis-associated gene 373 undergoing significantly low expression in
 PT subjects with high cedar pollen-specific immunoglobulin-E levels, useful
 PT in diagnosis of allergic diseases and screening drug candidates.
 XX Claim 2; Page 41-57; 80pp; Japanese.
 XX The invention relates to the human pollinosis-associated gene 373 which
 CC exhibits significantly reduced expression in the T-cells of individuals
 CC with high cedar pollen-specific IgE (immunoglobulin E) levels. The gene
 CC was isolated from T-cells from individuals allergic to cedar pollen using
 CC the differential display method. The invention also relates also relates
 CC to the protein encoded by pollinosis-associated gene 373; expression constructs and
 CC host cells comprising pollinosis-associated gene 373 nucleic acids;
 CC pollinosis-associated gene 373 primers and probes; antibodies against the
 CC protein encoded by the gene; methods of detection of pollinosis-
 CC associated gene 373 nucleic acids; and a method of diagnosis of allergic
 CC diseases via the detection of pollinosis-associated gene 373 nucleic
 CC acids. The invention additionally encompasses methods of screening drug
 CC candidates for the treatment of allergic disease by measuring the
 CC expression of pollinosis-associated gene 373 in pollen antigen-stimulated
 CC T-cells in the presence of a test compound relative to a control.
 CC Pollinosis-associated gene 373 is useful in the diagnosis of allergic
 CC diseases and in the screening of drug candidates for the treatment of
 CC such diseases. The present sequence represents human pollinosis-
 CC associated gene 373 cDNA
 XX
 SO Sequence 6245 BP; 1861 A; 1237 C; 1345 G; 1801 T; 0 U; 1 Other;
 Query Match 2.8%; Score 143.4; DB 3; Length 6245;
 Best Local Similarity 46.2%; Pred. No. 9.7e-30;
 Matches 769; Conservative 0; Mismatches 851; Indels 45; Gaps 7;

QY 485 GAAGCCCAAGTGGCCGAGAGTCTCTCAGACGATGTGAAGCGCTTCAATGAGTACCG 544
 DB 415 GAATTTTTCATCTGAGAAAGCGTGTGTCTGTCTTAATGAACCTTAAACATTTTGG 474
 QY 545 CACCAGCAGCTAGTGTCCGCTCTATATACCAACGATTTGTGCACTTCTCGAGATCC 604
 DB 475 GGTTCAGCATTATGATCAGACCAATTAATGATTCACATTTGTGTATTCGTTTGGAGACTG 534
 QY 605 CACAGTTAATGTGAGGAGGCGGCCATCAACCGTACTGAAATCTACAGCATGTAG 664
 DB 535 CAACAGTCAAGTGAAGATGCTGCAATTTGGCTATATGAGATTATTAAGCATGTGG 594
 QY 665 GATTCGATTGGCCCAAGCTCCGTCGATGAGCGATGTCTGCG-----TCGAATTTGG 720
 DB 595 AGAAAAATGTGAGATGATCTTTATTAAGAGAGAAATTCCTCGTCAATTAAGAAATGAT 654
 QY 721 CTATGTTGAGCAAAAGTTGACCAAGGTCAACAGAGAGGTCTACTGCTACCTTGAGCC 780
 DB 655 ATTTCGCAATTTGATGATGAGTGAAGTTCAGGCGGTATGATTTGAGTGTCTGCAAGA 714
 QY 781 TTAATAACAGAAATGCAATGAGTGAAGTGGCTTGACAGAGCCGACATATTGGGTTGAGG 840
 DB 725 TAAAGCTTCGATGATGAAGATCAGTGAATGAAATAGCGCATCATGATGATGATCAGC 774
 QY 841 AGCAGCCACAGATGATTTAAGCGGCACTACCTGGCCGTTTCG----- 887
 DB 775 CTTCAAGGTTCTTGACCTTAAACATCCGAAATCTTCCAACTGCAAGGAAAGCTTCG 834
 QY 888 -TCATCACTGCGCCCAAAACCAATGTGAAGATGACCGGTGATGCGCGCGCTAAC 946
 DB 835 TTCACAGATGGCCCTTAAGTTGAGAGTCTTCTTAAGAAAGAGGTGCGAGGAGTTCA 894
 QY 947 CATGAAATCTTTCGATCTTGAATCTTGAAGTGTGCGCAATTTGAACATCTTCCAGCTTA 1006
 DB 895 TGAAGATGATTTTAAAGCTTTTACAGATGCTCCCTTCAATTCGATTTATTTAGTCG 954
 QY 1007 GGACATGACGATATCTCAAGCAAGTACTGTATCATAGTGTAAAGCGAGACTG 1066
 DB 955 AGAACTTCAGAGAAATTAATTAATTAATGAAGAAATTTGTCAGATGATTAACATGACTG 1014
 QY 1067 GGAGAAAGCTGTGATGCTCTCAAGAAATGACAGGATTTGCTCATTTCACTATCAAC 1126
 DB 1015 GATATGAGTGTGCATATGACATGAGAAATTCAGATCACTGTCTTT-----GCTGGAGC 1068
 QY 1127 TCAAGCGAGATTGTGTGTGATACAGCTTAAGAAATTTGCTTAAGCTTGTGCAATCT 1186
 DB 1069 TGCAAGTATGATGCTTTTTCACATTTAGATTTGATGAGGACCTTAAACTTTC 1128
 QY 1187 CAAGGAGAACTAGCATCAAGTATTCGCGAGCGGTGATCAACATCGCCTAATGTC 1246
 DB 1129 AGCTAAGGATCTTAATCCAGGTGTGAGAACTTGTATTTACTGTAGCCACCTTTC 1188
 QY 1247 TAAAGCGTGAAGAAATTAATAGATGCTTCTGTGAGACATTTTGAACCTGATTTAA 1306
 DB 1189 AACAATTTGGAAACAAAGTTTATGATAGCGCTGAACCATTTATACCTACACTTTTAA 1248
 QY 1249 TCTGTCCCAATAGTGAAGAAAGTATGAGCAACTTCTGAGATGCGACATCAGATTAT 1308
 DB 1307 TTTAATAGAAACAGCGGGAAGTATGATCGGCTTCCACATTAAGCTTGAATAT 1366
 QY 1367 CATTAATATACATGACCAAGCTGCTCAAGATCTACAGACGCTGAAATCAATC 1426
 DB 1309 CATTGGGATATCTATGATACCAAGCTTATACCTTTAATACAGCAATTTGCA--CATC 1365
 QY 1427 AAAGTGAAGGATTAAGTCCACACTGTGTAGCTGATGTGCTCTTTCGAGAGTGG 1486
 DB 1366 AAATATAGTTCCTGTGAGAGAGAGCTTCAATTTAATTTATATTTATGTTTCAAGAGTGG 1425
 QY 1487 GCAGAGGAAGCGTTGAAAGGAATGCCACGTTAAGGACACCTTAAATAATTCAT 1546
 DB 1426 GCAGACTATTCATTTGAAAGCATGACGCGTCTGTGTTGAATTAATTAAGAGGAAT 1485
 QY 1547 TGGCATGCAAGCTGCGATGACGCGCATTTCCAGATTAACGCTTATGGGCTTTCAAGGG 1606

DB 1486 TCATATCTCTGCTGAGGCGGAGGTGAGGCAAGAAAGACATCATGGCTTTAGAA 1545
 QY 1607 TCATCTTCAGAGCTGGCGGATCAATATATGAAATATGACATTAAGCTGGCCAGCGGC 1666
 DB 1546 CCACTTCTCTGTGTAAGCTGAAACATTTATATATTCCTTGAGCCATCTTATCAGAAAG 1605
 QY 1667 ATTAGAAAGGAACAGAGGCGGCGGAGAGAGAACTGTATC 1711
 DB 1606 TCTTCAACTTACTTAAGAGATTCTGGCAGGTGTACATCTCTCC 1650

RESULT 5

ADB62363
 ID ADB62363 standard; cDNA; 2524 BP.

XX ADB62363;

XX 04-DEC-2003 (first entry)

XX Human cDNA encoding clone FCBBF30105860.

XX Human; ss: gene; pharmaceutical; diagnostic; gene therapy;
 tissue regeneration; cell regeneration; membrane protein;
 signal transduction-related protein; transcription-related protein;
 osteoporosis; neurological disease; cancer; tumour.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 355..2524
 FT /*tag=a
 FT /partial
 FT /product="Clone FCBBF30105860 protein"

XX EPI308459-A2.

XX 07-MAY-2003.

XX 28-MAR-2002; 2002EP-00007401.

XX 05-NOV-2001; 2001JP-00379298.

XX 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I,
 Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI: 2003-450961/43.
 DR P-PSDB; ADB64333.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
 marker or medicines for regulation of their expression and activity, or
 as targets of gene therapy.

XX Claim 1; Page: 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 from 1970 fully defined nucleotide sequences which encode novel
 polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 or its partial peptide, an antibody binding to the polypeptide or peptide
 of the polynucleotide, immunologically assaying the polypeptide or peptide
 with the antibody of the encoded protein, and observing the binding
 between the two, a transformant carrying the polynucleotide in an
 expressible manner and an antisense polynucleotide. The oligonucleotide
 is useful as a primer for synthesizing the polynucleotide, or as a probe
 for detecting the polynucleotide. The polynucleotides and encoded
 proteins are useful as pharmaceutical agents and many disease-related
 genes may be included in them, for developing a diagnostic marker or

CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a cDNA of the invention. Note: Some of the sequence
 CC data for this patent is not represented in the printed specification, but
 CC is based on sequence information supplied by the European Patent Office.
 XX

SQ Sequence 2524 BP; 650 A; 590 C; 671 G; 613 T; 0 U; 0 Other;

Query Match 2.8%; Score 141.8; DB 10; Length 2524;

Best Local Similarity 46.1%; Pred. No. 1.6e-29;

Matches 768; Conservative 0; Mismatches 852; Indels 45; Gaps 7;

QY 77 GAAGCCAGAGGACCTGATGCTTTATCCAGCAAAATCCCAAGCGGACATGCTGTGAA 136
 DB 357 GGAAGCCCCGAGCATGAGTACTTCTGCGCCAGGTGCAGCAAGAGACGTGGCGCG 416
 QY GGTACAGCTGCGCGAGGATCTGTGACATTCCTTAGCG---ACGACAAATCAATTTGT 193
 DB 417 GCTGAGAGTGGCCAGAGCTCTGCTTACCTTGGCGCCCGCCCATCTGAGACT 476
 QY 194 CTGACGAGCATGAGATTCTCT-----TATTGACGTTTGATGCGCATGCTGAGCGG 244
 DB 477 GAGAGAGGACCTGGGCGCGCTAGGCAAGACAGTGCAGCGCTCACCGGCTGGGTTC 536
 QY 245 CAGCACTTTAAGTTGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATAAGCGATT 304
 DB 537 GAGCACTTACCGGATATCAATTAATGAGTGAATTAAGTCCCTTGGGCAAGATT 596
 QY 305 GGGCAGGATTTTAAATGATACACAGGCTACGTTCTCCACATGTGATCGATCGCTGGG 364
 DB 597 ATCAACAGCTTTAAATCTATGAGCAATGTTATGTTAGCTTAATAGCAAAATGGG 656
 QY 365 AGACAGAGGAGACACAGTCCGAGAAAGCGCAACTTCTCTGCGGACCTCATGAGCA 424
 DB 657 AGATGCCAAAGACAGAGTTCGAGTGAAGCTCAGACTCTGATATTAAGTAAAGATCA 716
 QY 425 CAGAGTCTTCCGCGCGCGCTGATGACAGAGCTGGCCCATAGCTTCAAGCA 484
 DB 717 AGTAG-----CACCACTATGATCAATTTGGAGCAATTTGCTGTTTAAACCA 770
 QY 485 GAAAGCCAGAGTGGCGGAGGATTCCTTCAAGCATTTGTAAGCGCTTTCATGATGACG 544
 DB 771 GAATTTCCATCTCGAAGAGCGGTGTCTGTCTTATTAAGAACTTAACATTTTGG 830
 QY 545 CACCCAGACCTTAGTGTGCGCTTAATTAACCAACAGTTTGTGACCTTCTGGAGATCC 604
 DB 831 GGCTCAGCAGCTAGTCACTAGCAAAATGATACACATTTGTGTCTGTTGGAGACTC 890
 QY 605 CACAGTTAATGTAGAGAGCGGCCATCCAAAGCTAGTGAATTCACAAAGCATGTAG 664
 DB 891 CAACAGTCAGAGTAGAATCTGCAATATTTGGCTAATGTGAGATTAATGACATGTGG 950
 QY 665 GGATGATGCGCCAGACCTCCGTGCGATGAGCATGTTCCTGCGC---TCGAATTTGG 720
 DB 951 AGAAAAAGTAGATGATCTTTAATAGAGAGAAATTCCTGCTAGTTAGAAATGAT 1010
 QY 721 CTATGTTGAGCAAAAGTTGACCAAGTCAACAGAGGCTTAATCTCACTTACGCC 780
 DB 1011 ATTGGCAAAATTTGATTAAGTGAAGTTGAGGCGGTATATTTGAGTCTTCCAAAGA 1070
 QY 781 TTAATAAACAGATGGAATGGAATGAGTGGCTTGAACGAGCGCAATATTTGGTGAAG 840
 DB 1071 TAAAGCTTGAATGAGTAAGAAATCAGTGAATGGAATAGCCATCACTGATGATGAC 1130
 QY 841 AGGACCCACAGAGATGATTAAGCGGCACTACCTCGCGCTTTCG-----887
 DB 1131 CTTCAGGTTCTGACACTTAACATCCGGAATTCCTGCCAACAGTGCAGAAAGCTGG 1190


```

QY 1107 CTCATTCTGCTATACACCTGACCGCGAGTTTGTGCTGTAACGTAAGAATTGTCG 1166
DB 533 CTGTGTGCTGAG-----CTGCACAGTATGATGCTTTTTCACATTACGATTGTG 586
QY 1167 TTAAGCTTGAGACATCTCAAGAGAGAACTACATCAAGGTATCCGAGGCGAGC 1226
DB 587 GATGAGACCTTAACCTTACAGTAAGATCTTAGTCCAGGTGTTAGAGAACTGTG 646
QY 1227 ATACCATGCGCTTACATGCTCTAAGACGCTGAGAAATAAGATGCTTGTGAGAC 1286
DB 647 ATATCTGTAGCCACCTTTCAACAGTTTGGAAACAAGTTGATCATGGCGCTGAACC 706
QY 1287 ATTTTGAGACCTGATTTAATTATATACGAAACAGCGGCAAGGTATGATCCCTTCC 1346
DB 707 ATTGTACCTACACTTTTATCTCGTCCCAATAGTGCAAAAGTATGCAACTTCTGGA 766
QY 1347 ACAATAGCTCTGAGATATCATTAAGTATACATGACGACCAAGCTGCTCAAGTAC 1406
DB 767 TGTGACGACATCAGATTATCATCTGCGATCTCATGATCCAGACTTATACCTTTAATA 826
QY 1407 ACAGACACTCTGATCAATCAAGTCAAGAGACATAGGTCACACTGTGAGCTGATG 1466
DB 827 ACAACCAATTGCA-----CATCAATTCAGTCCCGTAGAGACGTTCAATTTGAATTTTA 883
QY 1467 GTGCTGCTCTTGAGAGAGTGCGAGACGAAAGCGTTGGAAGGATGCCACTTAAGG 1526
DB 884 GATTATATTGTGCAAGAGTGCGAGACTCATTTGGAAGACATGCAAGCGCTTGTGTT 943
QY 1527 GACACCTTAAAAAATTCATTTGGCGATGCGAGCTGCGAGCGCGCATTTCCAGATAC 1586
DB 944 GAAATATTAAAAAGGAATTCATATGCTGACGCTGAGCGCAGAGTGAGGCAAGAAAG 1003
QY 1587 GCGTATGCGCTTTCAGGCGCTCACTTTCAGAGCTGCGGATCAATATATGAACATTA 1646
DB 1004 ACATCATGAGGCTTATGAACAACATTTCTGTGTGAAGCTGAACATTTATATTTCCCTT 1063
QY 1647 GACATAGCTGCGGCGGCGCATTTAGAAAGGAGACGAGGCGGCGAGAGAGAACT 1706
DB 1064 GAGCCATCTTATCAGAGAGTCTTCAAACTTAAAGAGTTCTGGGAGTGTAGCATCT 1123
QY 1707 GGTAC 1711
DB 1124 CTTCC 1128

```

RESULT 7
ADB79874
ID ADB79874 standard; DNA; 5614 BP.

```

XX ADB79874;
XX AC
XX DT 04-DEC-2003 (first entry)
XX DE Rat putative KIAA0627 coding sequence. SEQ ID 114.
XX KM Analgesic; pain; streptozocin-induced diabetes; rat; gene; ds.
XX OS Rattus norvegicus.
XX PN EP1279744-A2.
XX 29-JAN-2003.
XX PD 26-JUL-2002; 2002EP-00255249.
XX PF 27-JUL-2001; 2001GB-00018354.
XX PR 07-FEB-2002; 2002GB-00002910.
XX (WARN) WARNER LAMBERT CO.
XX PI Brookbank RA, Dixon AK, Lee K, Plincock RD;
XX WPI; 2003-395407/38.

```

```

DR P-PSDB; ADB79873.
XX Use of isolated gene sequences and encoded polypeptides that are
XX upregulated in the spinal cord in response to streptozocin-induced
XX diabetes for screening compounds for the treatment of pain, or for
XX diagnosing pain.
PS Claim 1; Page 208-210; 334pp; English.
CC The present invention relates to nucleotide sequences which are useful in
CC the screening of compounds for the treatment of pain, or for the
CC diagnosis of pain. The nucleotide sequences are up-regulated in the
CC spinal cord in response to streptozocin-induced diabetes. The present
CC sequence is one such nucleotide sequence.
SQ Sequence 5614 BP; 1668 A; 1120 C; 1181 G; 1645 T; 0 U; 0 Other;

```

Query Match 2.3%; Score 116.6; DB 10; Length 5614;
Best Local Similarity 48.7%; Pred. No. 6.2e-22;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

QY 927 GGTATGCGCGCGCGCTAACCATGAAATCTTGAATCTAGCTTGAAGTGTCCGCA 986
DB 269 GGAAGTCTGAGACGATGATGAAGATGATTTATAAAGCTTTACAGATGTCCCTCT 328
QY 987 TTGAACATCTTCCACCGTAAAGACATGACGATATCTTCAAGCAAGTACTAGTATC 1046
DB 329 ATTCAATTTATTTCTAGTCAGAACTCGAAGAAACATTAATAAATCAGGGAATTTTG 388
QY 1047 AGTATTAAGCAAGCAGCTGAGAAACCTGTGATGCTTCAAGAAATCAGGCGATTG 1106
DB 389 TCAATATTAACATGATGCTGAGATCAGCGCAATGCAATGCAAGAAATTCATCTG 448
QY 1107 CTCATTCTGCTATACACCTGACCGCGAGTTTGTGCTGTAACGCTAAAGAAATTGTCG 1166
DB 449 CTGTGTGCTGAG-----CTGCACAGTATGATGCTTTTTCACATTTACATTTGTTG 502
QY 1167 TTAAGCTTGTGACATCTTCAAGAGAACTACGATCAAGTATCCGAGCGCTGC 1226
DB 503 GATGAGACACTTAACTTCAAGTATGATCTTAAATCCAGGCTGTAAAGAACTGTG 562
QY 1227 ATGACCATGCGCTTACATGCTTAAAGCGTGAGAAATTAATGATGCTTGTGTGAGC 1286
DB 563 ATTAGCTGACCCCACTTTCAGATTTTGGAAACAAGTTGATGATGCGCTGAAGCC 622
QY 1287 ATTTTGAGACCTGATTTAATTATATATACAGACGCGGAGGCTTGTGATCCGCTTCC 1346
DB 623 ATTTGACTTACACTTTTATATCTGCTCCCAATAGTCAAAAGTATGCACTTCTGGA 682
QY 1347 ACAATAGCTCTGAGATATCATTAAGTATACATGACACCAAGCTGTCAAGATCTAC 1406
DB 683 TGTGACGACATCAGATTTATCATTTCCGCAATCATGATACCCAGACTTATATCTTATA 742
QY 1407 ACAGACACTCTGATCAATCAATGAAGTGAAGACATTAAGTTCACACTGTGTAGCTATG 1466
DB 743 ACAAGCAATTGCA-----CATMAAATCATGTTCCCGTGAAGAGAGCTTATTTTGA 799
QY 1467 GTGCTGCTTTCAGAGAGTGAGAGACGAGCGCTTGAAGAGATGCGCACTAAGG 1526
DB 800 GATTATATTGTGCAAGATGAGAGCTCATTTCTTGAAGAGATGAGCGCTTGTGTT 859
QY 1527 GACACCTTAAAAAATTCATTTGCGATGACGCTGCGATGACCGCCCATTCAGATAC 1586
DB 860 GAAACTATTAAGGAATTCATGATGCTGACGCTGAGGCGCAGAGTGAGGCAAGAAAG 919
QY 1587 GCGTATTTGGCTTTCAGGCGCTCACTTTCGAGACTGCGGATGAATATATATGAACTTA 1646
DB 920 ACATACATGGGCTTGAAGAAACACTTTCGTGAAGCTGAACATTTATTAATTCCTT 979
QY 1647 GACATAGCTGCCGCGCATTAAGAAAGGAGACGAGGCGGCGAGAGAGAGAACT 1706
DB 980 GAGCCATCTTATGAGAAAGTCTTCAAACTTAAAGTTCTGGCAGTGTAGCATCT 1039

```


QY 1707 GGTAC 1711
 Db 1040 CTTCC 1044

RESULT 8
 ADX07257
 ID ADX07257 standard; DNA, 5614 BP.
 AC ADX07257;
 DT 21-APR-2005 (first entry)
 DE Cyclin-dependent kinase modulation biomarker DNA SEQ ID NO 1822.
 XX cytostatic; cyclin-dependent kinase; cdk; biomarker; gene; de.
 XX Homo sapiens.
 OS WO2005012875-A2.
 PN 10-FEB-2005.
 PD 29-JUL-2004; 2004WO-US024424.
 PF 29-JUL-2003; 2003US-0490890P.
 PR (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA Li M, Rupnow BA, Webster KR, Jackson DG, Wong TW;
 PI WPI; 2005-163068/17.
 DR P-PSDB; ADX07258.
 PT Biomarkers useful for predicting or determining the response of a mammal
 PT to a cancer treatment comprising administration of a modulator of cyclin-
 PT dependent kinase activity.
 PS Claim 5; SEQ ID NO 1822; 141pp; English.

CC This invention describes a novel method of predicting or determining
 CC whether a mammal will respond or is responding to an anti-cancer agent
 CC that modulates cyclin-dependent kinase (cdk) activity. The method
 CC comprises measuring the level of one or more biomarkers selected from
 CC 277 biomarkers given in the specification (nucleotide sequence SEQ ID
 CC NO:1246 (Genbank EST W28729) is especially preferred). The method of the
 CC invention is utilized in a kit for determining or predicting whether
 CC patient would be susceptible or resistant to treatment by an agent
 CC modulating cdk activity. The invention also describes a method for
 CC utilizing individualized genetic profiles for treating diseases and
 CC disorders based on patient's response and molecular level, specialized
 CC microarrays comprising the biomarkers described, antibodies directed
 CC against the biomarkers and a cell culture model to identify biomarkers.
 CC The cdk modulator is preferably N-5-[(1,1-dimethylethyl)-2-oxa-
 CC oxazolo[1,1-b]thio]-2-thiazolo[1,4-b]pyridine carboxamide, 0.5-L-
 CC tartaric acid salt. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pat_sequences. This
 CC sequence encodes a biomarker used in the method of the invention.

CC Sequence 5614 BP; 1668 A; 1120 C; 1181 G; 1645 T; 0 U; 0 Other;
 SQ

Query Match 2.3%; Score 116.6; DB:14; Length 5614;
 Best Local Similarity 48.7%; Pred. No. 6.2e-22;
 Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCCGCGCCCTAAACCATGAATCTTTGAACTAGCTTTAGAGTGTCGCCGAA 986
 Db 269 GGAAGTGCTGAGCACTGTGTAAGATGATTTATTAAGCTTTTACAGATGCCCTCT 328

QY 987 TTGAACATCTTCCACCGCTAAGACATGACGATATCTTACAGCAAGTACTAGTATCATC 1046
 Db 329 ATTCAATTTATTTCTAGTCGAGAACTCGAAGAAACATTAATAATCAAGGAAATTTTG 388

QY 1047 AGTGATAAAAACGACATCGGAGAAACGTGTGATGCTCTCAAGAAATCATGAGGCATTG 1106
 Db 389 TCAGATGATTAACATGATCGGATCAGCTGCCAATGACATGAAATTAATTCATCCTG 448

QY 1107 CTCATCTCAGCTATCAACATCCAGCCGAGTTGTGTGCTATACAGCTTAAAGAAATTGTG 1166
 Db 449 CTGTGTGTGAG-----CTGCACAGATGATGATGCTTTTTCACATTTTACATTTGTTG 502

QY 1167 TTAAGCTTCGTGACATCTCTCAAGAGAACTAGACATCAGATGATCCGAGGCGTGC 1226
 Db 503 GATGAGACATTAACCTTTCAGCTAAGATCTTATGATCCAGGTGTGTAAGAAAGCTTGT 562

QY 1227 ATCAACATCGCTTACATGCTTAAAGCGCTGAAGAAATTAACATGATCCCTTCTGCGAGC 1286
 Db 563 ATTACTGTAGCCCACTTTCAACAGTTTGGGAAACAAAGTTGATCATGGGCTGAAAGCC 622

QY 1287 ATTTTGAGACACGTATTAATTAATACAGAACGCGGAAAGTGCATGCAATCCGCTTC 1346
 Db 623 ATGTACCTACACTTTTAAATCTGCTCCCAATGTGCAAAAGTATGCAACTTCTGGA 682

QY 1347 ACAATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAAAGCTGCTCAAGATCTAC 1406
 Db 683 TGTGAGCAATCAAGTTTATCATTCGGCATATCATGATACCAACTTAACTTAAATA 742

QY 1407 ACAGACACTGTGAATCAATCAAGTCAAGAGACATAGCTCAACTGTGTAGCTGATG 1466
 Db 743 ACAAGCAATGTCA---CATCAAAATCAATGCTCCGAGAGACAGTTCATTTGAAATTTTAA 799

QY 1467 GTGCGCTCTTCGAGAGTGGCAGACGAAGCGTTGAAAGAAATGCCAGCTATTAAGG 1526
 Db 800 GATTTATTTGTGCAAGATGGACAGCTCATTTGAAAGACATGACAGCCGCTTTGGTT 859

QY 1527 GACACCTTAAAAAATCCATTTGGCGATGACAGATGCGATGACGCGGCATTCAGATAC 1586
 Db 860 GAACATTAATAAAGGGAATTCATGATGCTGACGCTGAGGCCAAGTGAAGCAAGAAAG 919

QY 1587 GCGTATTTGGGCTTTCAGCGCTCACTTTTCAAGAGCTGGCGGATCAAAATATATGAACATTA 1646
 Db 920 ACATACATGAGGCTTAGAAGACCACTTCTCTGTGAAGCTGAAGCAATTTATATATTCCTT 979

QY 1647 GACATAGCTGCCCGCGGCATTTGAAGGGAAGGAGGCGCGAGAGAGAACT 1706
 Db 980 GAGCCATCTTTATCAGAAAGATCTTCAAACTTATTAAGATTTCTGGCAGTATGATCTT 1039

QY 1707 GGTAC 1711
 Db 1040 CTTCC 1044

RESULT 9
 AA158443
 ID AA158443 standard; cDNA, 6445 BP.
 AC AA158443;
 DT 22-OCT-2001 (first entry)
 DE Human polynucleotide SEQ ID NO 646.

XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 XX peripheral nervous system; neuropathy; central nervous system; CNS;
 XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 XX leukemia; ss.
 XX Homo sapiens.
 OS WO200153312-A1.
 PN 26-JUL-2001.


```

Db      596 GATGAGCACTTAACTTTCAGCTAAGAGATCTTAAATCCAGATGTTAGAGAACTTGT 655
Oy      1227 ATACCATGCGCTTACATGCTTAAAGACGCTGAGAAAATAACTAGATGCTTCTGTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCACAGTTTGGGAAACAAGTTGATCATGCGCCCTGAAGC 715
Oy      1287 ATTTGGAGCACCCTGATTAATTAATATACGAAACAGCGCGAAGTCAATGATCCGCTTC 1346
Db      746 ATTGACCTACACTTTTATATCTCGTCCCAATAGTGCAAAAGTCATGCAACTTCTGGA 775
Oy      1347 ACAATAGCTCTGAATATCTTATAGTATACATGACCAACCTGCTCAAGATCTAC 1406
Db      776 TGTGAGCAATCAGATTATCATTTGGCAATTCAGTACCAAGACTTATACCTTTAATA 835
Oy      1407 ACAGACACTCTGATCAATCAAGTCAAGAGACATCAAGGTCACACTGTGTAGCTGATG 1466
Db      836 ACAAGCAATGCA---CATCAAAATCAGTTCCTGTAGAGAGACGTTCAATTTGATTTTA 892
Oy      1467 GTGCTGCTCTTGGAGAGTGGCAGACGAAAGCGTTGGAAAAGATGCCACTTAAGG 1526
Db      893 GATTATTTGTTGCAAGAGTGGCAGACTCATTCATTGAAAAGACATGACAGCCGCTTGGTT 952
Oy      1527 GACACTTAAAAAATTCATTGCGATGACATGCGATGACCGCCCATTCAGATAC 1586
Db      953 GAAACTATTAAAAAGGAATTCATGATGCTGACGCTGAGGCCAGATGGAGCAAGAAAG 1012
Oy      1587 GCCCTATTTGGGCTTTCAGCGGCTCACTTTCAGAGCTGGCGGATCAATATATGAACATTA 1646
Db      1013 ACATACATGGGCTTCTAGAAACCACTTCTCGTGAAGCTGAAACATTTATATTCCTT 1072
Oy      1647 GACATAGCTGCGCCAGCGGCATTTAGAAAAGGAAAGGAGGCGGAGAGAGAGAACT 1706
Db      1073 GAGCATCTTATCAGAAAGTCTTCAAACTTAAAGAGTTCTGGCAGATGATACATCT 1132
Oy      1707 GGTAC 1711
Db      1133 CTTC 1137

```

RESULT 12

AA158442 standard; cDNA, 6487 BP.

AA158442;

22-OCT-2001 (first entry)

Human polynucleotide SEQ ID NO 645.

Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokine; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss.

Homo sapiens.

WO200153312-A1.

26-JUL-2001.

26-DEC-2000; 2000MO-US034263.

23-DEC-1999; 99US-00471275.

21-JAN-2000; 2000US-0048725.

25-APR-2000; 2000US-0052317.

20-JUN-2000; 2000US-00598042.

19-JUL-2000; 2000US-00620312.

03-AUG-2000; 2000US-00653450.

14-SEP-2000; 2000US-00662191.

19-OCT-2000; 2000US-00693036.

```

PR      29-NOV-2000; 2000US-00727344.
XX
XX      (HYSE-) HYSEQ INC.
PA
XX      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang Z, Mehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI      Zhou P, Goodrich R, Dymnac RT;
XX
XX      WPI; 2001-442253/47.
DR      P-PSDB; AAM39286.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders such
PT      as central nervous system injuries.
XX
PS      Claim 1; SEQ ID NO 645; 10078DP; English.
XX
CC      The invention relates to human nucleic acids (AA157798-AA161369) and the
CC      encoded polypeptides (AAM8642-AAM42213) with noctropic.
CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      localised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: Immune system suppression,
CC      Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukemias and
CC      C.N.S disorders. Note: The sequence data for this patent did not form
CC      part of the printed specification
XX
SQ      Sequence 6487 BP; 1969 A; 1248 C; 1327 G; 1943 T; 0 U; 0 Other;
XX
Oy      927 GGTGATCCGCGCGCGTAAACGATGAACTTTGCAATCTAGCTTGAAGTGTCCCGCA 986
Db      362 GGAGCTGCTGAGAGAGTGTAGTGAAGATGATTTTAAAGCTTTTACAGATGTCCCTTCT 421
Oy      987 TTGAACATCTTCACGCTTAAGACATGAGACATCTCAAGCAAGTACTGATCATC 1046
Db      422 ATTCAGATTATTTTACTGACGACCTGGAAGAACTTAATTAATTAACAGGAAATTTTG 481
Oy      1047 AGTATTAACAGCAGCTGGAGAAACGTGTGATGCTCTCAAGAAAGATCAGGCAATTG 1106
Db      482 TCAGATATTAACATGACCTGGAGATCGGCTCCATGCTCATGAAATATTCATCTG 541
Oy      1107 CTCATCTCAGCTATCAGCTCAGCCGCGAGTTGTGCTGTACAGCTAAAGAAATTTG 1166
Db      542 CTGTGCTGTGAG-----CTGCACAGTATGATGCTTTTTCACATTTACGATTTGTTG 595
Oy      1167 TTAAGCTTGTGAGACATCTTCAGAGAGAACTAGATCAGAGATCCGAGAGCGTGC 1226
Db      596 GATGAGCACCCTTAACCTTCAAGCTTAAGATCTTAGATCCAGGGTGTAGGAAGCTTGT 655
Oy      1227 ATCAGCATGCGCTCATGTCTTAAGACGCTGAGAAATTAACATGATGCTTCTGTGAGC 1286
Db      656 ATTACTGTAGCCCACTTTCACAGTTTGGGAAACAAGTTGATCATGCGCCCTGAAGC 715
Oy      1287 ATTTGGAGCACCCTGATTAATTAATATACGAAACAGCGCGAAGTCAATGATCCGCTTC 1346
Db      746 ATTGACCTACACTTTTATATCTCGTCCCAATAGTGCAAAAGTCATGCAACTTCTGGA 775
Oy      1347 ACAATAGCTCTGAATATCTTATAGTATACATGACCAACCTGCTCAAGATCTAC 1406
Db      776 TGTGAGCAATCAGATTATCATTTGGCAATTCAGTACCAAGACTTATACCTTTAATA 835
Oy      1407 ACAGACACTCTGATCAATCAAGTCAAGAGACATCAAGGTCACACTGTGTAGCTGATG 1466
Db      836 ACAAGCAATGCA---CATCAAAATCAGTTCCTGTAGAGAGACGTTCAATTTGATTTTA 892

```

Query Match 2.3%; Score 116.6; DB 4; Length 6487;

Best Local Similarity 48.7%; Pred. No. 6.8e-22;

Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

DE Novel human cDNA SEQ ID NO 323.
 XX ss; cancer; neurodegenerative disease; human.
 KW Homo sapiens.
 OS US2003104529-A1.
 XX PN 05-JUN-2003.
 XX PD 04-JAN-2002; 2002US-00037270.
 XX PF 21-JAN-2000; 2000US-00488725.
 XX PR 25-APR-2000; 2000US-00552317.
 XX PR 19-JUL-2000; 2000US-00620312.
 XX PA (ZHOU/) ZHOU P.
 PA (TANG/) TANG Y T.
 PA (LIUC/) LIU C.
 PA (ASUNDI/) ASUNDI V.
 PA (DRMANAC/) DRMANAC R T.
 XX FI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
 XX DR WPI; 2003-678194/64.
 XX PT New polynucleotide, useful for treating diseases e.g., cancer or
 XX PT neurodegenerative diseases.
 XX PS Claim 1; SEQ ID NO 323; 99pp; English.
 XX CC The invention relates to a polynucleotide comprising a sequence given in
 CC the specification, or its mature protein-coding portion, or its
 CC complement. The polynucleotide is useful for treating diseases e.g.,
 CC cancer or neurodegenerative diseases and many others listed in the
 CC specification. The present sequence represents a novel human cDNA. Note:
 CC the sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?docid=2003104529.
 XX SQ Sequence 6487 BP; 1969 A; 1248 C; 1327 G; 1943 T; 0 U; 0 Other;
 Query Match 2.3%; Score 116.6; DB 9; Length 6487;
 Beet Local Similarity 48.7%; Pred. No. 6,8e-22;
 Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;
 QY 927 GGTGATGCCGGCGCCGTAACCATGGAATTTTCGATCTGAGTGTCCCGCAA 986
 DB 362 GGAGGTGCTGGAGCAGTTGATGAAGATGTTTATATAAACTTTACAGATGTCCTTCT 421
 QY 987 TTGAACATCTTCACAGCTAAGACATGAGATATCTACACAGTACTAGTATC 1046
 DB 422 ATTCGATTTATTTCTAGTGAAGAACTCGAAGAACTTAAATTAATCGGAAATTTTG 481
 QY 1047 AGTGATATAAAGCAGACACTGGAGAAACGTGTGATGCTCTCAGAGATCAGGCATG 1106
 DB 482 TCAGATGATATAACATGACTGGATCAGCTGCCAATGACATGAAGAAATTCGATCAGT 541
 QY 1107 CTCATTTTCAGCTATACACTCAGCCGCGATTGTGCGCTGACAGTAAAGAAATTTGCG 1166
 DB 542 CTTGTGTGCTGAG-----CTCCACAGTATGATGCTTTTTCACATTTTACGATTTGTTG 595
 QY 1167 TTAAGCTTGAGACATCTCTCAAGGAGAACTACGATCAGATGATCCGAGGCGTGC 1226
 DB 596 GATGAGACCTTAACCTTCACGCTAAGATCTTAATGATCCAGATGTTAGAGAACTTGT 655
 QY 1227 ATACCATGCGCTACATGCTTAAGACGCTGAGAAATTAACATGATGCTTCTGCTGAGC 1286
 DB 656 ATTACTGTAGCCCACTTTCACAGTTTGGGAAACAGTTTGCATGATGCGCTGAAGCC 715
 QY 1287 ATTTTGAGACACTGATTAATTAATTAACAGAACAGCGCAAGTCAATGCAATCCGCTTCC 1346
 DB 716 ATTGTAACCTACACTTTTAAATCTGTCGTCCTCCCAATAGTCAGAAAGTCAAGCAACTTCTGGA 775

QY 1347 ACAATAGCTCTGAAGTATATCATTAAGTATACATGCAATGCAAGTCTGCAAGATCTAC 1406
 DB 776 TGTGACAGAAATCAGATTTATCATTTCCGCATATCTCATGTACCCAGATTAATACCTTAAATA 835
 QY 1407 ACAGACACTCTGATTCATTAATGAAAGTCAAGACATTAAGTCTCAACACTGTGAGCTGATG 1466
 DB 836 ACAAGCAATGCA---CATCAAAATCAGTTCCCGTGAAGAGACCTTCATTTGAAATTTTAA 892
 QY 1467 GTGCTGCTCTTCGAGAGTGGCAGACAGAGCGCTTGGAAAAGTAATGCAACCTGACTAAGG 1526
 DB 893 GATTATTTGTTGCAAGATGGCAGACATCATTTCAATGGAAGACATGACGCCGCTTGTGTT 952
 QY 1527 GACACCTTTAAAAAATCATTTGCGATGACAGACTGCGATGCAATGCAAGCCGCAATTCAGATAC 1586
 DB 953 GAAACTATTAATAAAGGGAATTCATGATGCTGACCGCTGACGACAGAGTGAAGCAAGAAAG 1012
 QY 1587 GCCTATTGGGCTTTCAGGCGTCACTTCCAGAGCTGGCGGATCAATATATATGAACATTA 1646
 DB 1013 ACATACATGGGCTTTAGAAAACCACTTTCCTGTGAAGCTGAACATTAATATATTCCTT 1072
 QY 1647 GACATAGCTGCCAGCGCGCATTTAGAAAGGAAAGGAGGCGGAGAGAGAGAACT 1706
 DB 1073 GAGCCATCTTATTCAGAAAGAGCTTCAAACTTAATTAAGATTTCTGCGAGTATGCAATCT 1132
 QY 1707 GGTAC 1711
 DB 1133 CTTC 1137
 RESULT 15
 AA160228
 ID AA160228 standard; cDNA; 6491 BP.
 XX AA160228;
 XX 22-OCT-2001 (first entry)
 XX DE Human polynucleotide SEQ ID NO 4217.
 XX Human; noctropic; immunosuppressant; cytostratic; gene therapy; cancer;
 KW peripheral nervous system neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 XX OS Homo sapiens.
 XX PN WO200153312-A1.
 XX PD 26-JUL-2001.
 XX PF 26-DEC-2000; 2000WO-US034263.
 XX PR 23-DEC-1999; 99US-00471275.
 XX PR 21-JAN-2000; 2000US-00488725.
 XX PR 25-APR-2000; 2000US-00552317.
 XX PR 20-JUN-2000; 2000US-00598042.
 XX PR 19-JUL-2000; 2000US-00620312.
 XX PR 03-AUG-2000; 2000US-00653450.
 XX PR 14-SEP-2000; 2000US-00662191.
 XX PR 19-OCT-2000; 2000US-00693036.
 XX PR 29-NOV-2000; 2000US-00727344.
 XX PA (HYSF-) HYSF INC.
 XX PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
 PI Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 DR WPI; 2001-442253/47.
 DR P-PSDB; AAM41072.

XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 4217; 10078bp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AA157798-AA161369) with neurotrophic,
CC immunosuppressant and cytoskeletal activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX

SO Sequence 6491 BP; 1965 A; 1251 C; 1328 G; 1947 T; 0 U; 0 Other;

Query Match 2.3%; Score 116.6; DB 4; Length 6491;

Best Local Similarity 48.7%; Pred. No. 6.8e-22;

Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

OY 927 GGTGATGCCGGCCGCTTACCATGGAATCTTTCAGATCTAGCTTTGAGTGTCCCGCA 986
DB 362 GGAGGTGCTGGAGCGATTGATGATGATATTTTATTAAGCTTTTACAGATGTCCCTCT 421
OY 987 TTGAACATCTTCCAGCTAAGACATGAGCATCTACCAAGACATGACTAGTATCATC 1046
DB 422 ATTCGATTTTATTCTAGTCGAGACTCGAAGAACTTAATTAATCAGGAAATTTTG 481
OY 1047 AGTGATTAAGACGACACTGGAGAGAAAGCTGTGATGCTCTCAGAGATCAGGCGATG 1106
DB 482 TCAGATGATTAACATGATCGGATGATGATGATGATGATGATGATGATGATGATGATG 541
OY 1107 CTCATTTCTAGCTATCAGATCAGCGGCGATTTGCTGCTGATGATGATGATGATGATG 1166
DB 542 CTGTGCTGAGG-----CTGCACAGATGATGCTTTTTCACATTTAGATGTTG 595
OY 1167 TTAAGCTTGTGAGATCTCTCAAGAGAACTAGATCAGATGATGATGATGATGATGATG 1226
DB 596 GATGAGACATTAACCTTTCAGCTAAGATCTTATGATCCAGGTGTGAGAGACCTTGT 655
OY 1227 ATCAACATGCTTACATGCTTAAAGAGCTGAGAAATTAATAGATGCTTGTGAGAGC 1286
DB 656 ATTACTGAGCCACCTTTCACAGCTTTTGGAGAAAGTTGATCATGGCGCTGAAGCC 715
OY 1287 ATTTTGAGCAGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1346
DB 716 ATTTGACCTACCTTTTAAATCTCTCCCATAGTGCAGAAAGTATGCAACTTCTGGA 775
OY 1347 ACAATAGCTGAGATATATATCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1406
DB 776 TGTGAGCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 835
OY 1407 ACAGACATGCTGATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1466
DB 836 ACAACCATTTGCA---CATCAAAATCAGTTCCTCGAGAGAGAGCTTCAATTTGAATTTTA 892
OY 1467 GTGCTGCTTCTGAGAGAGTGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1526
DB 893 GATTATATGTTGCAAG 952
OY 1527 GACAGCTTAAATAATCAATGAG 1586
DB 953 GAAACTATTAATAAG 1012
OY 1587 GCCTATGAGGCTTTCAG 1646

DB 1013 ACATACATGGGCTCTTAAAGCACTTCTGTGTAAGCTGAAACATTATATATTCCTT 1072
OY 1647 GACATAGCTGCCAGCGCGCATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1706
DB 1073 GAGCATCTTATCAG 1132
OY 1707 GGTAC 1711
DB 1133 CTTC 1137

Search completed: November 29, 2005, 15:58:23
Job time : 1891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: November 23, 2005, 15:02:54 (without alignments)
2022.166 Million cell updates/sec

Title: US-10-030-850-2
Sequence: 1 MAYRRPSDLDGFIQMPKAD.....RNCISGGSSSTKNSSAASSS 1492

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 572060 seqs, 82675679 residues
Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/ptodata/1/1aa/5-COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/6-COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/H-COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/PCITUS-COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/RE-COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1236.5	16.2	255	US-09-270-767-32225	Sequence 32225, A
2	1236.5	16.2	255	US-09-270-767-47442	Sequence 47442, A
3	1214.5	16.0	916	US-10-104-047-2297	Sequence 2297, Ap
4	920	12.1	723	US-10-104-047-2487	Sequence 2487, Ap
5	502.5	6.6	454	US-10-104-047-2083	Sequence 2083, Ap
6	248	3.3	51	US-09-270-767-61363	Sequence 61363, A
7	248	3.3	66	US-09-270-767-45835	Sequence 45835, A
8	208.5	2.7	1025	US-09-248-796A-20573	Sequence 20573, A
9	196.5	2.6	2283	US-10-172-502-4	Sequence 4, Appl1
10	193.5	2.5	2137	US-09-134-001C-4463	Sequence 4463, Ap
11	191	2.4	4377	US-09-949-016-6978	Sequence 6978, Ap
12	182	2.4	3913	US-09-949-016-10933	Sequence 10933, A
13	173.5	2.3	2781	US-09-698-295-10	Sequence 10, Appl1
14	173.5	2.3	2907	US-09-698-295-11	Sequence 11, Appl1
15	172	2.3	2733	US-09-949-016-11433	Sequence 11433, A
16	172	2.3	3259	US-09-949-016-6507	Sequence 6507, Ap
17	171.5	2.3	1140	US-09-538-092-647	Sequence 647, Appl
18	171	2.2	907	US-08-938-830-26	Sequence 26, Appl
19	171	2.2	907	US-09-020-222-26	Sequence 26, Appl
20	171	2.2	907	US-09-068-377-26	Sequence 26, Appl
21	167	2.2	1237	US-09-862-027-78	Sequence 78, Appl
22	163.5	2.1	787	US-10-104-047-3340	Sequence 3340, Ap
23	162.5	2.1	686	US-09-252-991A-20509	Sequence 20509, A
24	162.5	2.1	2842	US-07-741-940-7	Sequence 7, Appl1
25	162.5	2.1	2842	US-08-289-548A-7	Sequence 7, Appl1
26	162.5	2.1	2842	US-08-452-654-7	Sequence 7, Appl1
27	162.5	2.1	2842	US-08-449-731-7	Sequence 7, Appl1

28	162.5	2.1	2843	US-08-452-655B-2	Sequence 2, Appl1
29	162.5	2.1	2843	US-08-452-655B-7	Sequence 7, Appl1
30	162.5	2.1	2843	US-08-450-582-2	Sequence 2, Appl1
31	162.5	2.1	2843	US-08-450-582-7	Sequence 7, Appl1
32	162.5	2.1	2843	US-09-538-092-1007	Sequence 1007, Ap
33	162.5	2.1	2973	US-08-821-355A-7	Sequence 7, Appl1
34	162.5	2.1	2973	US-09-003-687A-7	Sequence 7, Appl1
35	162.5	2.1	2973	US-09-136-605-7	Sequence 7, Appl1
36	161.5	2.1	313	US-09-252-991A-29545	Sequence 29545, A
37	161.5	2.1	2843	US-10-092-138A-30	Sequence 30, Appl
38	161.5	2.1	2843	US-08-681-219A-30	Sequence 30, Appl
39	161.5	2.1	3696	US-09-134-001C-5080	Sequence 5080, Ap
40	161	2.1	2663	US-09-538-092-1252	Sequence 1252, Ap
41	160.5	2.1	526	US-09-252-991A-28761	Sequence 28761, A
42	158	2.1	1401	US-09-976-594-1035	Sequence 1035, Ap
43	157.5	2.1	564	US-09-949-016-6898	Sequence 6898, Ap
44	157.5	2.1	564	US-10-069-540A-2	Sequence 2, Appl1
45	157	2.1	565	US-09-949-016-6902	Sequence 6902, Ap

ALIGNMENTS

RESULT 1
US-09-270-767-32225
Sequence 32225, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 32225
LENGTH: 255
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-32225

Query Match	Best Local Similarity	Score	DB 2;	Length	255;
Query Match	96.5%	Pred. No. 1.1e-95;	Matches 247;	Conservative 0;	Mismatches 8;
Indels 1;	Gaps 1;				
QY	859	LISLTQYADGKELTQOLKCVLDMFRKMFNDTHKTVSLFLDTVELILVHANEFSRNG	918		
DB	1	LISLTQYADGKELTQOLKCVLDMFRKMFNDTHKTVSLFLDTVELILVHANEFSRNG	59		
QY	919	SSSCITRLFNKLGTDLNSMHSKIWKTLQVHVEYPTLOLQKELFRITSDSTQTTTKR	978		
DB	60	LFIILTRFNKLGTDLNSMHSKIWKTLQVHVEYPTLOLQKELFRITSDSTQTTTKR	119		
QY	979	IAILRFLDLNNTCKSSDPSPDOSQACERTVYLKLAOLAADQKSELRSQARSCVALYN	1038		
DB	120	IAILRFLDLNNTCKSSDPSPDOSQACERTVYLKLAOLAADQKSELRSQARSCVALYN	179		
QY	1039	INTPMTLLADLPVYODSARSCISHMRROSOCSNGANSPSSSPKPLQSPS	1098		
DB	180	INTPMTLLADLPVYODSARSCISHMRROSOCSNGANSPSSSPKPLQSPS	239		
QY	1099	VGFPSLQSHHQSLI 1114			
DB	240	VGFPSLQSHHQSLI 255			
RESULT 2					
US-09-270-767-47442					
Sequence 47442, Application US/09270767					
Patent No. 6703491					
GENERAL INFORMATION:					
APPLICANT: Homburger et al.					
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster					

```

FILE REFERENCE: file Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17
NUMBER OF SEQ ID NOS: 62517
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 47442
LENGTH: 255
TYPE: PRT
ORGANISM: Drosophila melanogaster
US-09-270-767-47442

```

```

Query Match          16.2% Score 1236.5; DB 2; Length 255;
Best Local Similarity 96.5%; Pred. No. 1.1e-95;
Matches 247; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

```

```

QY 859 LISTLOVADGKELTQOOLKCVLDFMRKMPDTHTKVYSLFDVTELLVHANEISNG 918
DB 1 LISTLOVADGKELTQOOLKCVLDFMRKMPDTHTKVYSLFDVTELLVHANEISNG 59
QY 919 SSSCLTFLFNKLGTDLNSMHSKIWKTLQVHVEYFPTOLKELFRILISDSTQPTTKTR 978
DB 60 LFLITRLFNKLGTDLNSMHSKIWKTLQVHVEYFPTOLKELFRILISDSTQPTTKTR 119
QY 979 TAILRFLTDLANTYCKSSDFPSDSQACERFVYLKLAQADQKSWELRSQARSCVALYN 1038
DB 120 TAILRFLTDLANTYCKSSDFPSDSQACERFVYLKLAQADQKSWELRSQARSCVALYN 179
QY 1039 LNTQOMTLLADLPVYODSARSCHSHMRROSOSCNAGSPSSPLSSSPKPLQSPS 1098
DB 180 LNTQOMTLLADLPVYODSARSCHSHMRROSOSCNAGSPSSPLSSSPKPLQSPS 239
QY 1099 VGPFASLQSHHQLSI 1114
DB 240 VGPFASLQSHHQLSI 255

```

```

RESULT 3
US-10-104-047-2297
Sequence 2297, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241e1 full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2297
LENGTH: 916
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-2297

```

```

Query Match          16.0% Score 1214.5; DB 2; Length 916;
Best Local Similarity 32.8%; Pred. No. 7.3e-93;
Matches 321; Conservative 180; Mismatches 322; Indels 157; Gaps 29;

```

```

QY 9 LDGFIQMPKADMRKVQLADLVTLSDDTNSIVC---TDMGFLIDGL-MPWLTGSHF 63
DB 5 MESCLAQVLQKDVQKRLQVQGLIDYFSDKQKSALEHDMQMLDLKVLATSWNSNY 64
QY 64 KIAQKSLAEFSELKRLGSDNATATATLPHVIRLGRSDPTVEKQULLRDMERHVL 123
DB 65 KVVLLGMDIISALVTRLQDRFKAQIGIVLPSLIDRLGAKOSVREDDQFTLLKTIWD-QAA 123
QY 124 PPOALIDPLATSCFHKNAKAREEFLQTIIVNALEHYGTQOLSVRYIIPVYALGDPVTN 183
DB 124 NPQVYVMDRLGG-FRHKNFRTREGICLCIATIANASGQITLISKIVYHICNLGDPDSQ 182
QY 184 VREAIQTLVEIYKHVGDRLRDLRRMDVPASKLAMLEQKFDQYKQSGLLPALSAXTN 243

```

```

DB 183 VRDAIINSLVEIYRHVGERVRADLSK-KGLPQSLNVIYFKFPEVQKSGMWIQA----- 236
QY 244 GNGVGLDEADNIGLRSEPTMIRKPLHSAVSSSLRPKPNVDVYG----- 288
DB 237 -NDRKPFDESDVD-GNRPS-----SASSTSKAPPSSRRNVGMQTTTRRLSGSTLQSK 286
QY 289 -----DAGAVTMSPESSFEVYVPOLNIFRAKQMDIYKQVLIISDKNADMERKVDALK 342
DB 287 SSAKAGAGAVDEEDFIKARDVVPVQIYSSRLDEESINKIRIELISDQKDWEOQVNAALK 346
QY 343 KIRALLILSYHTQOPQAVOLKELSLSFVUILKEELRSQVIRPACTIATAMSKTLRKLD 402
DB 347 KIRSLLAGAAYEDNFQO-HLRLLDGA-F-KLSAKDLRSQVREACITLGHLSVLAGKFD 404
QY 403 AFGSILIEHLINLIONSAXVIAASTALKYIIKYTHAPKLLKIYPTDLNOSKSDIRST 462
DB 405 HGAELIPTFNLIPNSAKIMATSGVAVVAILIRHTIIPRLIPYITSNCT-SKSVAVRR 463
QY 463 LCEIMVLLFEEWQTKALERNATVLRDTLTKSICDADGADARRSHRYAVAFRRHPELADQ 522
DB 464 CFEFLDLLQEWQTHSLERHISVLAETIKKGIHDADSEARIEAKCYWGFHSFSREAEH 523
QY 523 IYGTLDIAAQRALEREREKGGGGGTGCTGTATPRTTVERIGTPTLOKPTPSMSIS 582
DB 524 LYHTLESYKALQSHLK-----NSDSIVSLPQSDRSS---SSQSLMRPLSAKISPT 574
QY 583 AVDTAAQRAKVAQVLYRSORRKLPGPNNSNOASMTGAASGLPFRPRLNSNGCTP-- 640
DB 575 GSTTSRASTVSTSVSTTGSLSR-----SRDIDVNAALSA---KSXVSSSGTTPFS 624
QY 641 ---ATPGRSV-----TPPRGRAG---V 657
DB 625 SAALPPGYSYASLESRRHREDMEYIGDSGRIRTRRQSSGATVASTPNNRGSRRAKV 684
QY 658 SGOPOGSRSTSPETKLRDQVGI-GNYTRG-ATGAIIPKXSGIPRSTASRETSPTSGG 715
DB 685 SGOPOGSRSSPGKLSGGYGLTGSSRPVTPSSSEKSKLPRSGCCRETSPTNIG- 743
QY 716 GLMKRSYSTGAGS-RTTFRNNPVRPSAPARLACGREAEHTLGVDDQOPDYVSGDYM 774
DB 744 -----LDRFGLOPGRIIPGSVANM-----KVLSTDELAHV-----ADAL 779
QY 775 RSGGMRGRKLMGRDESDIDISASVCSERSFDSYTRGNKSNYSLSGHTRLDWSTQR 834
DB 780 EK-PYRRRYEPYGVYSDDANSDASSVCSERSYGS--RNGGIPHY-----L 822
QY 835 APFDITITIOFCASTWSEKRGGLISLTOYLDGKELTQOOLKCVLDFMRKMPDTHTK 894
DB 823 RQTEDEVAEVLNHCASSNWSERKGLGLQWLKSORLTSVELKRLCEIFTRMFADPHSK 882
QY 895 -VYSLFDVTELLVHANE 913
DB 883 RVFSMFLVTLVDFIIHKD 902

```

```

RESULT 4
US-10-104-047-2487
Sequence 2487, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241e1 full length cDNA
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2487
LENGTH: 723
TYPE: PRT

```



```

; Sequence 45835, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 45835
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45835

```

```

Query Match          3.3%; Score 248; DB 2; Length 66;
Best Local Similarity 98.0%; Pred. No. 2.4e-13;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Cy 697 GIPSTASSRETSPTRSGGLMKRSMYSTGASRRTPRRNNVPPSAPAR 746
Db 1 GIPSTASSRETSPTRSGGLMKRSMYSTGASRRTPRRNNVPPSAPAR 50

```

```

RESULT 8
US-09-248-796A-20573
; Sequence 20573, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20573
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (9), (11)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unkno
US-09-248-796A-20573

```

```

Query Match          2.7%; Score 208.5; DB 2; Length 1025;
Best Local Similarity 20.1%; Pred. No. 6.1e-08;
Matches 250; Conservative 208; Mismatches 429; Indels 357; Gaps 64;

```

```

Cy 329 DKNADMEKRVDAKKIRALILSYTOPOFAVOLKELSLFVDILKEELSOVIREACI 388
Db 38 ETEBNMKREKNILOMRKILNGNSATOPRSELOVOCITITANGCKGASSLTITLSSNCQ 97
Cy 389 TIAYMSKTLRNKLDFAFCMSILEHLINL--IONSARKVIASATLAKYIYTHAPKLLK 445
Db 98 LKSCAVILKKSLEFVASLFTPLIKLCSSTKNISTANMSVALVANYLPYT--SKMIQ 155
Cy 446 IYTLTANOSKSDIRS-TLCELMVLLFE-----EMQTKALERNAVTVRLDKLSIG 495
Db 156 RIT-LASDRNYQPPSYSLIWLHLLIKIGIDRSYIGHDSSFLEAANKVMKLLK---- 210
Cy 496 DADCDARRSHRYAVYAFRRHPELADQIYGTLDIAAORALR-REEGGGGGGTGTGTGTA 554
Db 211 DANENVQTAKECYWCFTFRVPEDEARLKLLEPIYVALERSQRESG-----GSGIA 263

```

```

Cy 555 PETRRIVSRIGRTPTGCT-PSMKRISAVDTAAQAVYRAQYTLXSROKRLGPNNS 613
Db 264 P--IRTLSS-----SRPSRLKE-----ALLENK-----ELRRRPPRSNSG 299
Cy 614 NQASMTGAASGSLPRPLRNSNSGTPATPGVTPPRPGAGVS--OSOPSRSTS---P 669
Db 300 EOST--KIKSVLPFRP-----TKSSRLKSLRPDVGKHSQPAVRASWTYP 345
Cy 670 STKLRDQYGGIGNYRGATGAIPIKASGIPRSTASRETSPTSGGLMKRSMYSTGAGS 729
Db 346 STQ-----SG-PKATPKQERSGTE----- 364
Cy 730 RRTPERNNVPPSAPAPALAQSRHAETLVGDGQPDYVS--GDYMRSGMGMKRM 786
Db 365 ---VHKSPLEISRPSRL--DTGAVSPFNKKDPMINFLSSSDSLIKGILNLYKAI 419
Cy 787 GRDESDDIDSEASVCSERFSDSYTGKNSYLSGSH-----TRLDMSTORAPF- 837
Db 420 GK--ENLPSRLSLK-----SISKHYQFMKPLTSDYTKKAAAL 460
Cy 838 ---DIIETIIQPCASTHWSERKGLISLTQYLADGKELTQOOLKCVLDMFRKMDHTK 894
Db 461 LRPD---FLRVCAIV-FDEPDEAVISLI-----ICCI-----DVLT- 493
Cy 895 VSIPLDTVELLIVHANETSRRNGSSCLTRLEN-KLG-TDLNMSKTIKTIQVYHEV 952
Db 494 ---FYESACNLTLVADTPNIGSHALVMOISNOKLITKLI-----LOALLIASKH 543
Cy 953 PPLOLKELEPRIISDSQTPTTKTRIALRPLTLANTYCKSDSPSDSQACERTVLK 1012
Db 544 AVTDVQGEIYFQEI-----VKLIIVKATDIYSLIC----- 574
Cy 1013 LAQLAADQKMELESRQASCLVALYNLTPQMTLLADLPKYVQDSARSC--IHSNMRQ 1070
Db 575 -----QLFRQ-----LYTIDSNKFLSLVEDVGKKEVEVFIIGISTVLE 616
Cy 1071 SGCNSGANSPPSSSPSSPKPL-QSPSVGPFPASLQSHHOLSTSPSRSSSVQOE 1129
Db 617 RPSSKLDYDYLTVKFTNLIGTFVLSQKPSADDFMLPKRSEFLSGSEVYNSKXVAKIEA 676
Cy 1130 LTPS---SELDIOHNOKTSEERHCFGGQYOTALAPNGNGLVHDOGOQDSCLSS 1185
Db 677 LDSPQKQSQMD-QISLKRNSB-----HQBSEVLDPE-ANVSITADG- 717
Cy 1186 SNSKQSSANTTOSNTESATMLDNL-----BERTONAKSPTD- 1227
Db 718 -GSKLNSKN--DDDNLKQEMERIDPLKPISNKIRKISYEDAKQNEKPRLENNWGQFQ 774
Cy 1228 ---AKYITVSIINAENGELI-----LASNLMESEVVRVALTTKQPPVELLQTS- 1273
Db 775 YAKFSRAIRVNA-VAENMSILTSQEFESCSKLVETPQSALIKVT--WFDVLSLVSSVDY 831
Cy 1274 ---LTNIGICIKGNCGLPKHKFRSIRMLNLILEAHTVDVYAGLHVLSKMSNMKMRH 1330
Db 832 QEEFLNG-----KHLKESLWEPFSKIDKLDS--LVNMGFLKLOLLFN--- 876
Cy 1331 NMNHFLELLIKLIIQCYQSHKEALRIDSMIPRIAPSLPDLG-----INIVNPIV 1381
Db 877 DTPRVDDLFPALVDTGQ-----EELDSLEYFIWMNMLSLNDELKMSPEKVALN--Y 928
Cy 1382 ATGEFFPNLCALIKLLE-----VTEHNGSEITDAH-LDIYFPNLABADDTQSVYKRAAV 1435
Db 929 LBEQ-ENNLITLISCLNLYLAKVYVDDCDLVAKIYRLDTFGKLFH---BREVFRKSAT 984
Cy 1436 FCIVKLYFVGEKVKPKL-SVLPSPK-----VRLNLYIIEK 1471
Db 985 ICYSNL---LAKTNVSEVVDITDKVKSRYPATYORLIEFTMR 1025

```

```

RESULT 9
US-10-172-502-4
; Sequence 4, Application US/10172502
; Patent No. 6841154

```

```

; GENERAL INFORMATION:
; APPLICANT: FOSTER, Timothy et al.
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
; FILE REFERENCE: P07263US01/BAS
; CURRENT APPLICATION NUMBER: US/10/172,502
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/298,098
; PRIOR FILING DATE: 2001-06-15
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 4
; LENGTH: 2283
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-10-172-502-4

Query Match      2.6%; Score 196.5; DB 2; Length 2283;
Best Local Similarity 18.3%; Pred. No. 2,5e-06;
Matches 171; Conservative 149; Mismatches 443; Indels 173; Gaps 25;

447 YTDLTNLSKSS-----KQIRSTLCELMVLLFEEMQTKALERNAVTLRPTLKKSIGDA 497
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1119 FSDSTSDSKASSTASSESIQSASSTSGSVSTSLSTSNSEKSTGMSDSTSLSTSES 1178
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      498 DCDARRHS-----RYAYWAFRRHPELADQIYGLDIAAQR-----LERERGGGGG 545
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1179 DSISESTSDSISASISESTFISLES-NSTSDSESQASAFISELSESTSESTSE 1237
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      546 GTGTGTGTAPETRRVRSIGRTPTGLQKPTPSMRISAVDRAAQRKVRQ---YTLXS 602
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1238 SVSSSTSESTSLSDSTSESGSTSTSLNSTSGSTISSTSLSESTFKSESVSTLSM 1297
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      603 RQRKPLGNNNSQASM-----TGAAAGSLPRPR 631
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1298 STSTSLSTSLSTSLSTSDSKSDSLSTSMSTSDSLSTSKSDSLSTSLSGSTSESE 1357
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      632 LNSNGGTPATTPGCVTPRPRGRAGVSGSQPGSRSTSPSTKLRDQYGGIGNYRGATGAI 691
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1358 SDSSTSESKSDSTSM-----ISMSQSTSGSTSTSTSLSDS----- 1396
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      692 PKKASGIRSTRASARETPTPSGGGLMKRSMYSTAGSRRTPERNNPVRPAPALLAQS 751
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1397 -----TSTSLSLASAMNOSGVDSNSASQASNSTSTSTSESOQSTSTSTGQSTQS 1448
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      752 REAHTLGVGDGCGDYVSGDYMRSGKMRGKLMGRDESDPIDEASVCSERFDSY 811
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1449 ESTSTSTSLSDST---ISKSTQSGSVSTASLSGSESEBSOSQSTSTASSTSEAST 1505
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      812 TRGNKSNYSLSG-----SHRLDMST---GRAPFDIET--ITQF 846
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1506 SLSDSTSTNSGASASTSTSLNSASASEDLASTSLSDSTASMSQSSSDQSTASISLD 1565
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      847 CASTHMERKQGLISLTQYLADGKELTQOQKLCVLDMPKXMDPHTKYSLFDVTVEL 906
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1566 SLSTSTSRKMTSTIASLSTSVSTSESGSTSESTSESDSTSLSDQST----- 1613
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      907 ILVHNENSRNGSSCLRLFNKLTGDLNLSMHSKI---WKTLOVVEHYFTQLOKEL 962
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1614 ---SRSTASGASTSTSTSDSRSTASTSTSMSTSTSDSQMSLSTSTSLSMDSSTSL 1669
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      963 FRISDSTQTP---TKTRIALRLFTDLANTYCKSDPFS---DQQA CERTVILKLAQL 1016
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1670 SDSVSDSTSDSTASATSGMSVSIISLSDSTSTSTASSEVMSASISDSQMSSEVNDSESV 1729
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1017 AADQKMEIRSQARCLVALYNLNPQMTLLADLPKYQDSARCSHSHMRQSCNS 1076
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1730 SESNSESPSKMSGSGSTVS--DSGSLSVSTSLRKSESVSESSSLC-----SQMSD 1779
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1077 GANSSSPFS-----SSSPKPLQSPVSPFASLQ---SHHQLSISTSPRQSSVEQ 1128
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1780 SVSTSDSSLSLSTSLRSESVSESDSLSDSTSGSTSTSTSGSLSTSTSLSGSESVSE 1839
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      1129 ELLFSESLDIQHNIOKT--SEIRHCGFGQYOTALAPNGFNGLQYHDQGDQSCASLSGN 1187
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCU
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO: 4463
; LENGTH: 2137
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-4463

Query Match      2.5%; Score 193.5; DB 2; Length 2137;
Best Local Similarity 19.0%; Pred. No. 4e-06;
Matches 162; Conservative 123; Mismatches 426; Indels 141; Gaps 20;

476 TKALERNATVLRDTLKKSIGDADCARHRSRYAYWAFRRHPELADQIYGLDIAAQRAL 535
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1301 TSLSEISTSVSDSTASSTSDSASTSESE-----SDASSTSLSGSTSTSL 1347
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      536 EREREGGGGCGTGTGTAPETRRVRSIGRTPTGLQKPTPSMRISAVDRAAQRKVR 595
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1348 SDSSTSTSDASASTSTSESESRASTSLSGSTSLSDSTSTSDSASTSTSV----- 1401
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      596 AQTLYRQRKPLGPNNSNOASMTGAAGSLPRPRLNSNGGTPATTPGCVTPRPRGRA 655
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1402 -----SDSNAS---TSLSGSLSTSVSDSTSTSDSASTSTSDSASTSESDSERA 1443
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      656 GVSQSGPGRST---SPSTKLARDQY---GIGNYRGATGAI PKKASGIRSTRASARET 708
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1444 STSL--GSTSTSLSDSTSTSTSDSASTSTSVSESNSTSTSLSESLSTSVSDSTSTSD 1501
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      709 SPTSGGGLMKRSMYSTGAGSRRTPERNNPVRPAPALLAQS REAHTLGVGDGCGDY 768
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1502 SASSTSTSVSDSASTSSSESVSTSDSESTSTSDSASTSTSVSESNSTSTSLSGSTST 1561
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      769 VSGDYMRSGKMRGKLMGRDESDPI-----DSEASVCSERFDSYTRGN 815
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1562 SVSSTSTSTSDASASTSESDSDSASTSSSESVSTSVSDSTSTASSTSTSVSDSN 1621
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      816 KSNYSLSGS--HTRLDMSTORAPFDIETITQFCASTHMERKQGLISLTQYLADGKELQ 874
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1622 SASSTSLSTSTSLSDSTSMSTSDSAST---STSESDSDSASTSLD----- 1665
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      875 QQLKCVLDMPKXMDPHTKYSLFDVTVELILVHNENSRNGSSCLRLFNKLTGDL 934
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1666 -----STSTSV--SESTSTSTSTSVASNSTSTSLSDSTSTSLSDSTST-- 1707
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      935 LNSMHSKIWKTLQVVEHYFTQLOKELFRISDSTQPTTKTRIALRLFTDLANTYCK 994
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1708 -STSESGSTSTSESDSDSASTSLS--ESTSTSLSDSTSTST-----SDSASTSM 1754
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```



```

QY 1216 -----RTTQAKSPDADAKVITVS---INMAENGELLASINMESEVVALTLT 1262
DB 2106 DTILESPPDFSGHDDKSPFLSDSGFETRSKTPSPAQSAETTPKPLPHEVPIPVITET 2165
QY 1263 KQAPPELLQTSLTNIGICIKGNCNCLPN-----KHFPS 1295
DB 2166 KREVVHVRISYDPSAG-----DVPQTOPEEPVSPKSPFTFMELEPPTTSIKKVK 2217
QY 1296 IMRMNLINLEAHTVIVAGLHVLSKIMSNGRHNHMFLELILKIIQCYOHSKEALR 1355
DB 2218 AFQMASSBEDHN-----RVLSKGMVKKEETH-----ITTTBMVYHSPPGGE 2261
QY 1356 DIDSMPRIAPSLPLDLISINIVPAJGEPPTNLCAKILEVTEHHGSEITTDALDIY 1415
DB 2262 GASERIEE-----TMSVHDIMKAFQSGRDS-----KELAGLEPHKSA-----V 2300
QY 1416 PNLARSADDDQSMRKAAVFCIVKLYFVLGEEKVKPKLTVLNPCKRLLNYIEK 1471
DB 2301 SPDVHKSAAETSAQHAER-----DNQMKPKLE-----RIIEVHIEK 2336

RESULT 12
US-09-949-016-10933
; Sequence 10933, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10933
; LENGTH: 3913
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10933

Query Match 2.4%; Score 182; DB 2; Length 3913;
Beet Local Similarity 18.1%; Pred. No. 0.00011;
Matches 304; Conservative 241; Mismatches 613; Indels 518; Gaps 77;

QY 69 SLEAFSELKRLGSDPENATYATVLPVHIDRLGDSRDTVREKAQLLDLMEHRVLPQAL 128
DB 442 SIGANASASIRSSSD-RSIT-----LNRSSYARDSMWIE-ELVPSKQGLTFTREED 492
QY 129 IDKL-----ATSCFKAKN---AKVREEFLQT-IVNAL--HEXVQOOLSVRVIYP----- 172
DB 493 SDSLHYSWAADTLDLVNVLVSSPIHSGFLVSPFVVDARGSGMRGSRHGRKRIILPRKCTA 552
QY 173 -----VCAL-----LGDPTNVR-EAALQTLVEIKYKVGRLRLPDLRMDV----- 214
DB 553 PTRITCRVLKRXHKLANPPRMVGEGLASRLVEMGPAQFLGPVI---VEIHPFGSMRGK 609
QY 215 ASKLML-----EOKFDQVKGEGLLPSALKNTNGNGVGLDEADNIGLREPTRM 265
DB 610 ERELIVLRSENGETKEHOFDS-KNEDL-----TELLNGMDELDSPBELG-KKRICRI 662
QY 266 KR-PLHSAVSSSLRKPKN-VNDVTGAGAVTMESESSF-----EVPQ 307
DB 663 TNDPQYFVAVBRIGESNOIGPEGIILSTTVPLVOASFPEGALTKRIRVGLQAPVP- 721
QY 308 LNFIAKMDDIYKQVL-----VIISDKAADMEKRV----- 338

```

```

DB 722 -----DEIVKILGNKATFSPIVTVBERRRKFKHPIMTIIPVPPSGGVNGYXG 772
QY 339 DALKIRALLILSHTOQPOFAVOLKELS-----ISFV-----DIL 374
DB 773 DTPNRLRLCSITGTSIP-----AQMEDITGTPTPLTFIDCVSFTTNVSARFNLADCHOL 828
QY 375 KE-ELRSQVIREACITIAVMSK-----TLNKKLDAPCMS--ILEHLINLIONSAY 422
DB 829 ETVGIALQULRE-LICVPYNAKFFVFAKMPANDPVSSSLRCFMTDIXDKVKTLEQGENFEV 887
QY 423 IASASTALK-----YIKYTHAPKLLK-----IYDTLNOQ-----SKSDIRSTCEL 466
DB 888 ARSADIELEBKPIYVOCYGNLAPLTGGOQLVNFVFSFKENRILPFSIKIRIDTSOEPGR 947
QY 467 MVLFEEMQTALEARNATV-LRDTL-----KSIDADCDARHNR--YAWAFRRHPE 518
DB 948 LSPFKEPRTTGLQOTVANCNITLPAHKETESDODELEKTDRROSFALALKRYSY 1007
QY 519 LADQIYGLDIAQAALEREREGGGGCGTGTGTAPETRRVSRIGRTPGLOKPPPSM 578
DB 1008 LTF-----PMIERSTCAT 1021
QY 579 RSISAVDTAAQRAKVAQYTLVSRORRKLPPNNSNQ-----SMGAAASGSLPRRL 632
DB 1022 RSLP-----TYS-XKPFSTRPYQSMWTAPITVPPAKSGF---TSL 1060
QY 633 NSNGGTPATP-----GSVTPRP-----RGAGVYSGSPGSRSPSTKLDQYCGI--- 680
DB 1061 SSSSNTPSASPLKSIVSVSTPSPIKSTLGASTTSVSVISDVASPIRSFTMSSPIKTV 1120
QY 661 -----GNVYR-GATGAI PKKAGIPISTASRSTPSRSGGGLMKSMYSTG 726
DB 1121 VSQEPYNIQVSSGLARAPVATEATPLK--GLASNSTSRTSPYTTGSLIERSST-- 1176
QY 727 AGSRTPERNRPVPSAPARLAAQSREAHNTLVGDGQDPYVSGDYNRSGMRGRKLM 786
DB 1177 -----MTP-----PAPKSNIMYSSSLPRK-----SIITSAAPLISPLK 1212
QY 787 G-----RDESDDIISSEASVCSERSFDSSTYRGKNSYLSGSHRRLMSQRAFFDIE 841
DB 1213 SVSPVSKAVDVISAKITVMASSLSPPVKQMPGHAVALVNGSISLPIKYPSS-----S 1265
QY 842 TIIOFCASTHWSERKDLISLTQYLDGKELTQOOLKCVLDMFRMFMDDTHTKVYSLFD 901
DB 1266 TLINGCKAT-----ATLOEKISAKTNSVSVSAATVKEVFETTTAMPSPPLRS 1316
QY 902 TVTELLIVHANETSRNGSSCLTRLENKLDGLNLSMHSKIMKTQLVHHYEFPTLOLKE 961
DB 1317 YVS--AAPSAGSLRTPASA---LYTSLGSS-ISATTSSVTSSIIIVPVSVVAVLDEP 1370
QY 962 LFRISDSOTPTTKTRIAILRFLTDLA-----NTYCKSSDPSPSQSQCERTVUKLQ 1016
DB 1371 ALKKLIPDENS--FTKSAALILSPIKTLTTEHPQPHFERTSSPVKSLFLAPSAKLSTP 1428
QY 1017 AADQSMELRQANS-----LVALYNLNTPOMTLLADLP-----KVY 1055
DB 1429 SSLSSQELIKDVAEMKEDLMKMTNIILOTDVPEEKFPPELPKBERIDDEBFKIVEKVK 1488
QY 1056 QDSAR-----SCISHMRQOSQSCNSGANSPESSPLSSSPKPLQSPVGPASLQS 1107
DB 1489 EDLVAVSEILKKDVCVDNKGSPKSPKDKG--HSPEDMWIEFSSSEIRBA----- 1536
QY 1108 HHHQLSISTSPRQRQSSVEQELFSSGLDQHNIOKTSSEIRHRCFGQOYQALAPNGFN 1167
DB 1537 --RQQAASQSP-SLPERVQYAKAKASEKD--YNLTQYIDVLTNDIGSSSLT----- 1583
QY 1168 GHLOY-HDQGOQDSCASLSNSKTQSSANTQSNTPESATWRDLDERE----- 1215
DB 1584 -NLKTYFEADAKKGEERKRVLKPAIALQEKLMKP-PASRTTSSEKELCMQADSFRCT 1641
QY 1216 -----RTTQAKSPDADAKVITVS---INMAENGELLASINMESEVVALTLT 1262
DB 1642 DTILESPPDFSGHDDKSPFLSDSGFETRSKTPSPAQSAETTPKPLPHEVPIPVITET 1701

```

```

Qy 1263 KDOPEVLLQTSJLNLGICIKGNCELPLN-----KHFRS 1295
Db 1702 RREVVHVRISYDPSAG-----DVPQTPQPEBPVSPKSPFPMELPRTTSTJKEKXK 1753
Qy 1236 IIRMLNLINAEHTDVVLAAGLVLSKINSNQRNNHMFLELILIKIIOCYQSHKEALR 1355
Db 1754 AFQMKASSEEDHN-----RVLSKGMKVEKEETH-----YTTTTRMYVHSPPGGE 1797
Qy 1356 DIDSMPRIAPSLPLDLISININPVATGEPTNLCATKIILEVTEHHGSEITDAHLDIV 1415
Db 1798 GASENIEF-----TMSVHIDMKAFQSGRDS-----KELAGLEHNSA-----V 1836
Qy 1416 FPNLARSADDTQMYRKAAPFCIVKLYFVLGEEKYKPKLVNPSKVRLLNLYIEK 1471
Db 1837 SPDVHKSAAETSAQNAEK-----DNQMKPKLE-----RIEVIHIEK 1872

RESULT 13
US-09-698-295-10
; Sequence 10, Application US/09698295
; Patent No. 6689584
; GENERAL INFORMATION:
; APPLICANT: Jones, Michael H.
; TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY FACTOR
; FILE REFERENCE: 06501-068001
; CURRENT APPLICATION NUMBER: US/09/698,295
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: JAPAN 10/137631
; PRIOR FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10
; LENGTH: 2781
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-698-295-10

Query Match 2.3%; Score 173.5; DB 2: Length 2781;
Best Local Similarity 18.2%; Pred. No. 0.00031;
Matches 275; Conservative 189; Mismatches 487; Indels 559; Gaps 68;

Qy 5 KPSPDIDGFIQWPKADMRKVKQVLAEDLVTFISDDTNSIVCTDMGFLIDGLMPLTGSHPK 64
Db 1053 KVSDLASAGQETFKSKTKANDFFIDSKLASADDITGLICRKK-----KPLIOESBDT 1105
Qy 65 IAKSLAEAFSELIKRLGSDFNAYTATVLPVHVID-----RLG---DSRDVREKAQLL-LRDL 117
Db 1106 IVSSSKSLAHSSVPRSTNDRD---ATPLSRAMDFSGKGCDSSENSITLNSSDTVSIDS 1162
Qy 118 MEHRVLPQALIDKLTATSCFHKHAKAKVREELQITVNALHEGYTQQLSVRYVYIPVCAAL 177
Db 1163 SEEDMI-----VQNSNESISSEOF-----RTREOVVEVLEPKCELY 1198
Qy 178 -GDPTVNRBEAIIQTLVEIYKHVGRLL-----RP-DRRMDVVPASGLAMLEQKF 225
Db 1199 SGEISGENCE-----DRLPVKTEANGKPKSOOKKLEBRPVAKCS----- 1237
Qy 226 DQVKEGELLPSALNKT-----NGNGV-----GLDEADNIGLNERPTRMK 266
Db 1238 DQIK-----LAKTTDKNNENRESEKKGQRTSTQJINGKDKPKIYLKGCELEIS 1288
Qy 267 RPLHSAVSSSLRPK-PNNDVTGD-----AGAVTMSFESSFEVVPQL 308
Db 1289 E--SRVVGAVPEPKVNNINKIIPENDIKSLTYKESAIRFPIINGDVIMEDFERNSSETKS 1346
Qy 309 NIFAKMDMDIKQVLLVIL-----SDKN-----ADMEKKVADLKKIR 345
Db 1347 HLLSSDAEGNRYDLETLPTKESDSTOTTPPASCPESNSVNOVEDMEIETSEVKV- 1405
Qy 346 ALLILSYHTQPOFVAVOLKELSLSPVD-----ILKEELRSQVIREACT---TIA 391

```

```

Db 1406 -----TSSPITSEESNLSNDFIDENGLPINKNNVNGESRKRYITTEVTWTSTVA 1457
Qy 392 YMSKTL--RNLDAFCMSLIEHLINLIONSAKVASIALAKYI-----IK 436
Db 1458 TESTKIVKEGDK-----QTVSSTECNCKSVTTTTTIVTKLSTSTGSDVIDISVK 1511
Qy 437 YTHAPKLIKITYDTL-----NOSKXDIRSTCELMVLF-----E 472
Db 1512 EQSKTVATTVTTVDLTTGTLVTSMTVSKY-STRDKVKLMKFSRPKKTRSGTALPSYR 1570
Qy 473 EMQKALERNATV-LRTLKKSIGDADCDARHRRY-----AYAF 512
Db 1571 KFWKSTKKSIFVLPNDLKKARKGGIREVYPYNNAKPALDIWPPSPPTGITWRY 1630
Qy 513 RRRHPE-----LADQIYGLT--DIAQALALREREGGCGGTGTGTAPET----- 557
Db 1631 RLQTVKSLAGVSLMLRLMLASLRMDWAA-----KVPBGSGSTRITETSETITTEII 1683
Qy 558 -----RRVSRIGRTPGTLOKPTPSMRKSISA----- 583
Db 1684 KRDPVPGIRFVYCIRKIICPIGVPEPKETPTPKRGLASALRPKRPEPKQTCGVI 1743
Qy 584 VDT-----AAQRAKVAQVTLYGRQKRPPLGPNNSNOM 618
Db 1744 IETVAEELMELEIRAFERVEKEKAQAVEQAKKRL-----QOKPVIATST-TSP 1796
Qy 619 TGAASGSLPRRLNSNGGTPATTPGSVPRPRGAVSOSQSRSTSTKLKRDQYG 678
Db 1797 TSTSTTSTSPQKV-----MWAPISQSVT--TGTKVLTTKVGSPATVTFQOKNKHQ 1847
Qy 679 GIGNYR-GATGA-----IPKASG-IPRSTASRET-----SEPRSG-GGLM 718
Db 1848 TFAVWVQGGNSGVQVQKVLGITIPSTGTSOOTFTSPQRTATVTRINTSGSGGT 1907
Qy 719 KRSMTSTGASR-----RTERNNP-----VRSPARBLAQSREAHITLVG 761
Db 1908 SNSQVITGPQIRPMTVIRTPLOQSTLGKAIIRTPVWQPAPOVMTQIIR----- 1959
Qy 762 DDGQPDVYSGDYMSGGKRMGRKLMGRDESDPIDSEASVCSERSPSSYTRGNKSNYL 821
Db 1960 -GQP-----VSTAVSANITYSS-TPGOKSLTSA 1985
Qy 822 SGRHRLDWSFORAPFDIETIIIOFCASTHSEKDKGLISLTQYADGKELTQ-----QQ 876
Db 1986 T-STNSIQSSASQPP-----RPOQGVKLT--MAQLTQLTQHGNGQG 2025
Qy 877 LKCYLDMFRKKFMDTHTKVTSLFLDYTTELIVHANETSRNGSSCLTRLNKLGTDLN 936
Db 2026 LTVVIQCGGQ-----TTGQLQLIPQGVTVLPGPQ----- 2055
Qy 937 SMHSKIKKTLQVNAHEPFTQQLKELFRILISDSQPTTKRIRALRLFTLDMANTYCKS 996
Db 2056 -----OLMAMANGTVOFLFTPLATTITTSIT- 2086
Qy 997 DFPDSOQACERTVLIKLAOLADOKSMELRSQARSCLVALYNLTPQWTLILLADLPKYQ 1056
Db 2087 -----TYVSTTAAGGEQROSKL-----SPQW-----QVHQ 2112
Qy 1057 DS-----ARSCIHSMRQOSQNSGAN-SPSSSLSSSPKPLQSPSVGPAPSLQSH-- 1108
Db 2113 DKTLPPAOSSSVGPAAQOPQAPQSPAPQOPQOTQSPAPQOPQOTQSPAPQOPQOTQSP 2172
Qy 1109 -----HHQLGISSTSPASR-QSSVGEOLPSSSELDIHNIO-KTSEEIRHCFGGQVOTAL 1161
Db 2173 EAQPTHAOSKPYVAASQPOSNOQOSPVAVQSPQTRIRPSTPSQLSPEQOSQOVQTT 2232
Qy 1162 A-PNGFNCH--LQYHDQGOQDSCASLSNSKTQSSANT-----TQSNTPESA 1205
Db 2233 SQPIPIQPHSTLQIPSGQOPQOPQOVQSTQTLSSGQTLNOVSVSPSRPQLQIQOPQO 2292
Qy 1206 TMRLDNIRE 1215
; : : : :

```

Db 2293 VIAVPOLOOQ 2302

RESULT 14
US-09-698-295-1
Sequence 1, Application US/09698295
Patent No. 6689584
GENERAL INFORMATION:
APPLICANT: Jones, Michael H.
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY FACTOR
FILE REFERENCE: 06501-068001
CURRENT FILING DATE: 2000-10-27
PRIOR APPLICATION NUMBER: PCT/JP99/02340
PRIOR FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: JAPAN 10/137631
PRIOR FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 2907
TYPE: PRT
ORGANISM: Homo sapiens
US-09-698-295-1

Query Match 2.3%; Score 173.5; DB 2; Length 2907;
Best Local Similarity 18.2%; Pred. No. 0.00033;
Matches 275; Conservative 189; Mismatches 487; Indels 559; Gaps 68;

5 KPSDLGFIQMPKADKRVKQVLAEDLVFLSDDTNSICTDMEFLDGLMPWLTGSHFK 64
1179 KVSPLASGQEPKSKTKGNDFIDSKLASADDTGLTICKNK-----KPLIOESDT 1231
65 IAGSLKFSLSLIRKLGSDFNATVTVLPHVID---RLG---DSRDTREKAQLL-LBDL 117
1232 IVSSSKSLHSSVPEKSTYDNRD--ATPLSRAMDEFGKLGCDSESSSTLSSSDTVSIDS 1288
118 MEHRLPQALIDKLATSCFKHKNAKVREEFLOTTVNAHGYQQLSVRYIIPVCALL 177
1289 SEEDMI-----VQNSNESISEQF-----RRDEQVEVLEPLKCELV 1324
178 -GDPTVNVREAIOQLVEIYKHVGRLL-----RP-DLRMDVPAKMLMEQKF 225
1325 SGESTGNC-----DRLPVKGTGANGKKPSQOKLEERPVVKCS----- 1363
226 DOVQOEGLLPSALKNT-----NGNGV-----GLDADNIGLRERTRMIX 266
1364 DQIR-----LKNITTDKNNENRESEKKGQRTSTFOINGKDKPKIYLGCECLKEIS 1414
267 RPLHSAVSSSLRPK-PNVNDVTGD-----AGAVTMESFESSFEVFPQL 308
1415 E-SRVVSGNVEPKNNINIKIIPENDISLYKESAIRPFGINDVIMDFERNSSFTKS 1472
309 NIFAKDMDDIYKOVLVII-----SDKN-----ADMEKRVDAKKIR 345
1473 HLLSSDAEGNVRDLETLPTSKESDSTQTTSPASCPESNSVNOVDEMEIETSEVKV- 1531
346 ALLISTYTOFOFVAVOLKELSLFDV-----ILKEELSOVIREACI---TIA 391
1532 -----TSSEPTTSEESNLSNDFIDENGLPIKNENNVNGSKKKTIVITEVTTMTSTIVA 1583
392 YMSKTL--RNKLDACMSILEHLINLIONSAKVIASTIALKYI-----IK 436
1584 TESKTVIVKEKDK-----QIVVSTENCKASTVTTTTTIVTKLSTSTGSSVDIISVK 1637
437 YTHAPKLKIYDTL-----NQSASKDIRSTLCELMVLFF-----E 472
1638 EOSKTVVTTVTDLSLTGGTLVTMTVSKY-STRDKVLMKMFSPKTKTSGTALPSYR 1696
473 EMOTKALERNATVL-RDLTKSIGADADCDARHSKY-----AYNAF 512
1697 KFTVSTKTSIFVLBNDDKLKLARKGIREVPYFNNAKPAIDWIPYSPRPTFGITWRY 1756

QY 513 RRRPPE-----LADQIYGLT---DIAQRALEREGGGGGGTGTGTGAPET----- 557
Db 1757 RLQTVKSLAGVSLMLRLMLSLRMDWA-----KVPPGGSTRTETSETETITTEI 1809
QY 558 -----RRTVSRIGRPTGLQKPTPSMRSISA----- 583
Db 1810 KRQDVGYGIRFEXYCIKIKICPIGVPEPTPKETPTTPQKGLRSSALRKRPETPKOTGPVI 1869
QY 584 VDT-----AAQRAKVAQYTLVSRQKPLGAPNNNSQASM 618
Db 1870 IETVABEELWEIRAFARVEKEKAQAVEQAKKLF-----QCKPVIATST-TSP 1922
QY 619 TGAAGSGLRPRPLNSGSGTPATTPSSYTPRRGRAGVQSOSQSGSSSTSESTLRRQYG 678
Db 1923 TSSSTSTISPAOKV-----MVAPISGSVT--TGTRVLTTKVGSBATYTFQONKNPHQ 1973
QY 679 GIGNYR-GATGA---IPKASG-IPRSYASSRET-----SPTRSQ-GLM 718
Db 1974 TPATWVQOGSNGSVVOYQKVLGIIIPSTGTSQQTTSFQPRATVYIRNTSGSGGT 2033
QY 719 KRSMTSTGASR-----RTPERNNP-----VRSPARLLAQSREAHHTLGVG 761
Db 2034 SNGQVITGPQIRPGMTVIRTPLOOSTLGKALIRTPVWVGAPQOAVMTQILIR----- 2085
QY 762 DDGQPDYVSGDYMRSGGMRKMLMGDESDDIDSEASSVCSERSFOSSTYRGNKSNYSL 821
Db 2086 -GGP-----VSTVVASAPNTVSS--TPGQKSLTSA 2111
QY 822 SGRSTRLDWSTQRAPFDITIIIOFCASTHWSERKGLISLYGLADKEIYQ-----QQ 876
Db 2112 T-STSNIOSSASQPP-----RPOQGVKLT--MAQLTQLTOHGNGOG 2151
QY 877 LKCVLDMFRKMFMDTHTKVYSLFLDVTLEILVHANETSRNGSSCLTRLFNKLGTDLLN 936
Db 2152 LTVVIGQGGQ-----TGQQLDIPQGVTVLPBPGQ----- 2181
QY 937 SMHSKIKWTIQVHAIEPTPOLQELFRIIISDSQOTTTTKRIALRFLIDLANTYCKSS 996
Db 2182 -----OLMOAMPNGTVOFRFLTPPLATTAATTSITT----- 2212
QY 997 DFPESDQACERTYLAQALADOKSMELRSQARSCLVALYNLTPOMTLLADLPKYQ 1056
Db 2213 -----TVSTTAAGTEGROSKL-----SPQN-----QVHQ 2238
QY 1057 DS-----ARSCISHMRQSQSCNSGAN-SPSSSPLSSSPKLPQSPVGPASLOSH-- 1108
Db 2239 DKTLPRAGSSVGPAXKQPOPAQPSARPPQTPQPSAPQEPVQTOPEVQTOQTVSSHVS 2298
QY 1109 -----HHQLSISTSPSR-OSSVQELFLSSSLDIGNHQ-KTSEELRHCFGGQYQYAL 1161
Db 2299 EAQPTHAQSSKPPQVAAOSQPSQSNVQOSPVYRVQSPQTRIRPSTPQSLSPQOSQOVQTT 2358
QY 1162 A-PNGFRGH--IQVHDGQGDSCASLSNKGTSQANT-----TQSNTPESA 1205
Db 2359 SQPIPIQPHTSGLQPSQGPQSPQPOVOSTLSSGGTLNADVSSPSRPPQLIQGPQPO 2418
QY 1206 TMRLDNLERE 1215
Db 2419 VIAVPOLOOQ 2428

RESULT 15
US-09-949-016-11433
Sequence 11433, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: Venter, J. Craig et al.
TITLE OF INVENTION: POLYNORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: C0001307
CURRENT FILING DATE: 2000-04-14
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20
 PRIOR APPLICATION NUMBER: 60/237,768
 PRIOR FILING DATE: 2000-10-03
 PRIOR APPLICATION NUMBER: 60/231,498
 PRIOR FILING DATE: 2000-09-08
 NUMBER OF SEQ ID NOS: 207012
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO 11433
 LENGTH: 2733
 TYPE: PR
 ORGANISM: Human
 US-09-949-016-11433

Query Match 2.3%; Score 172; DB 2; Length 2733;
 Best Local Similarity 17.3%; Pred. No. 0.0004;
 Matches 270; Conservative 245; Mismatches 562; Indels 488; Gaps 64;

QY 71 EAFSELIRLGSDDFAATVATVPLPHVIDRLGDSRDVREKAOULLDLMEHRVLPQALID 130
 DB 562 EGFOLVYKWM-----NOTLDQKTNOI--DLQAEISENOAIHQ 597
 QY 131 KLATS-----CFKHKAKVREEFLOTIVVALHEGYTOOLS 165
 DB 598 KLITNTDASDSDSVALVETVISPCTGSSEHMKPELEKIL-----AL-EKEKEDQ 651
 QY 166 VRVYIPVYCALLGDPVTNVREAIOQLTVIYKHVGDRLRPDLRMDVPAKSLAMLEOKF 225
 DB 652 KKL-----QEALTSRKAILKKAQEKERH-----LREELKQCKD-----DYNLEQEF 693
 QY 226 DOVKOEGILLPSALK-----NTNGNVGDEADNT-----GLAERPRMTRK----- 267
 DB 694 DEOSKENENIGOLQLOIÖVESIDGKLPTDOESCSSTPGLPEPFKATEQHHTQPV 753
 QY 268 -----PLHSVSSSLPRKPNVDVYG-----DAGAVTME-----SFFSS 301
 DB 754 LESNLCPRMPSHSEDASALOGSTVAQIKAOKEIEAEKVELKAVSSSTSELTKESEV 813
 QY 302 FEVVPOLN-----IFHAKMDDIYKÖVLY 325
 DB 814 FOLQOINKQGLEISLKTVSHEAEVHAESLOOKLESSOLOIAGLEHRELPKDELOK 873
 QY 326 IISDKRAD-----WEKRVDAKIKIRAL-----LILSYHQPOFVAVQ-----LKEJSL 368
 DB 874 LISKEEDVSYLSGOLSEKEAALTKIQTEIIEQEDLIKALHTQLOMQAENHERIKOQV 933
 QY 369 SFVDI-----LKEELRSQVIREACTITAVMS-----KTLRKLDACFMSIEHIN 414
 DB 934 ELCEMKÖKPEEIGESRAKQOIQKLOALLISRKALKENKSLQEBL-SLANGTIERLTK 992
 QY 415 LI-----ONSAKVIASTIALKYIIKYTHAPKLIKITYDTPLNOSKSDIRSTLCE 465
 DB 993 SLADVESQVSAQNKKEK-----DTVIGRLALLQEBRDKLITEMDRSLLENQOS---LSSSE 1044
 QY 466 LMTLIFE-----EWOT--KALERNATVLRDITLKKSIGAD 498
 DB 1045 SLKTLLEGITEDKEKLVKEIESLKSKIAESTEMOKHEKLOKEYEIILQSYENSVNEAE 1104
 QY 499 C-----DARRHSRYAYMAFR-----RHPPELADOIYGLDIAORAL 535
 DB 1105 RIQHVVEAVRÖKÖBELYKGLRSTANKKETEKOLOEADQEMEMKEKAKKPKSKÖOKIL 1164
 QY 536 EREREGG-----GGGGTG-----TGCTAPETRRTVSRIGRTPGLOKPTPSMSISA 583
 DB 1165 ELEENDRLRAEVHPAGDTAKCEMETLLSSNASMKELERVMEYETLSKKFQSLMSEKD 1224
 QY 584 VDTAAQAKVAQÖTTLVSRÖKKPLGPNNSNOASMTGAAASGLPRPLNSNGSGTPATT 643
 DB 1225 SLSEEVODLKQIQEGVNSKÖANLEATEKHQDNQNTVY-EEGTOSIP-----GETEOD 1275
 QY 644 PGSVTPRRGRAGVSSQSGSRSTSPS--TKLRDQYIGNYRGATGAI PKKASGIPRST 702
 DB 1276 SLISMSTRPT-----CSESVPASASANPAVSKDFSSHDEINNYLQ-QIDQLEKIEIAGLEBEK 1330

QY 703 ASRETPSTBSGGIMKRSKMTSTGAGSRKTRPERNNPVRPAPARLLAQSREAEHTLGVCD 762
 DB 1331 QKNKEFSQTLNENKNTLLSOTISTDGLKMLQEB-----VTRKMLNQOIQEELS----- 1380
 QY 763 DGQPDVYGVYMRSGKMRGR-KLMGRDESDIDSEASSVCSERSFDSSTYTRGNKNSYL 821
 DB 1381 -----RVYKLKETAEBEKDDLEERLMLQALF-----L 1407
 QY 822 SGSHTRLDWSTORAPFDDIETIIQFCASTHWSERKDGILSTLOYIADQKELTQOQKCV- 880
 DB 1408 NGS-----IGNYQDVPDAQIKNELLE-----SEKKNLK-----KCVS 1440
 QY 881 -LDMFRKMFMDTRKQVSLFLDVTTELILVHANETSRNGSSCITRLE-----NKLGT 932
 DB 1441 ELEBEKQÖLVKERTKVESEIRKEVLEIKOGAKERGNKSHAKELQEBLKEKQEVKQÖLOK 1500
 QY 933 DLNLSMH--SKIWTLLQVHEYPPTQÖLQKEL-----FRITS 967
 DB 1501 DCIRYÖEKISALEETVAKAL-EFVQTBQO-KDLEITKENLAQAVEHRKKAQAEIASFKYLL 1558
 QY 968 DSTQPTTKTRIALIRFLDTLANTYCKSDPDSQACERTVYKLAQGLAADQKMEIRS 1027
 DB 1559 DDTQSEARVYADNLKXKKELOS--NKESVYSGMKQKQEDLERLEQ--AEKYLKEKK 1613
 QY 1028 QABSCVYAL-----YNLNTPQMTLLADLPKYVQDSASACIHSHRQSGSCNS--- 1076
 DB 1614 NMOEKTLALREKXVLEETIGEIQVTLNKXQ-----KEVQÖLOENLIDSTVT 1659
 QY 1077 --GANSFSSPLSSSPKQLOSPS--VGPPA-SLOSHHQLSI-----SSTSPRSQSSV 1126
 DB 1660 QLAAPFTKSMSSLODRDRVIDEAKWKEKRSDAIQSEEBEIRLQEDNSVYKQDQLOMSI 1719
 QY 1127 EQELLFSSSELDIQHNQ-----KTSEIRHCFGGQYOTALAPNGFNGLQYHDQ 1176
 DB 1720 HMEELKINISRLHDKQIWESKAQTEVOQCKVCDTLQGENKELLSQLEETRHL-YH--S 1776
 QY 1177 QQDSGASLSSNSKQSSANTQSTPESATRLNLE-----RETTONASPTDDAV- 1230
 DB 1777 SÖNBLAKLESLEKSKLDQDLSNLSLEKCEQÖKNLGGIIRQÖBADIQNSKFSYEQLETD 1836
 QY 1231 -----ITVSINNAENGELILASNLMESEVYVALTLTKQPVELLQTSLTNLGIC 1280
 DB 1837 LQASRELTSLRHEINKEKQKISLSLG--KEEALQYAL----- 1873
 QY 1281 IKGNCELPNKGFRSINKMLNILEAHTDVYIAGLVLSKIMSNOKRHNMMHFLLEIL 1340
 DB 1874 -----AELRQÖHDEKI--KELENLLSQEBEENIVLEBENKKAVDKTNQL-----METLKTIK 1923
 QY 1341 LKTIQ 1345
 DB 1924 KENIQ 1928

Search completed: November 23, 2005, 15:13:24
 Job time : 70 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 23, 2005, 15:06:50 ; Search time 10 Seconds

(without alignments)
452.221 Million cell updates/sec

Title: US-10-030-850-2

Perfect score: 7611
Sequence: 1 MAYRRPSDLGFTIQMPKAD.....RNCISGGSGSTKNSSAASSS 1492

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 17545 seqs, 3030971 residues

Total number of hits satisfying chosen parameters: 17545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_New:*

1: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTI_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the local score distribution.

SUMMARIES

No.	Score	Query Match	Length	ID	Description
1	191	2.5	4384	1	US-10-821-234-1120
2	186.5	2.5	761	1	US-10-485-517-252
3	163.5	2.1	1448	1	US-10-485-517-212
4	153.5	2.0	1402	1	US-10-971-982-2
5	146.5	1.9	3144	7	US-11-055-035-1
6	146	1.9	5024	1	US-10-793-626-2964
7	133	1.7	594	1	US-10-131-826A-10
8	131.5	1.7	482	1	US-10-821-234-972
9	125	1.6	1463	1	US-10-971-982-3
10	124	1.6	1618	1	US-10-984-645-2
11	123	1.6	1451	7	US-11-046-346-1
12	122	1.6	1189	7	US-11-074-176-134
13	121.5	1.6	1290	1	US-10-485-517-141
14	119	1.6	2897	1	US-10-499-715-2
15	116.5	1.5	1377	1	US-10-821-234-1070
16	114	1.5	1493	7	US-11-004-057-4
17	110	1.4	1279	1	US-10-793-626-3188
18	108.5	1.4	636	1	US-10-485-517-170
19	108.5	1.4	1637	1	US-10-821-234-1204
20	108	1.4	1125	1	US-10-821-234-1444
21	108	1.4	1493	7	US-11-004-057-21
22	108	1.4	2053	7	US-11-013-759-9
23	106.5	1.4	313	1	US-10-512-184-72
24	106.5	1.4	576	1	US-10-512-184-65
25	106.5	1.4	625	1	US-10-512-184-47

26	106.5	1.4	1207	1	US-10-821-234-1109	Sequence 1109, Ap
27	105.5	1.4	748	1	US-10-821-234-888	Sequence 888, App
28	105.5	1.4	989	1	US-10-821-234-975	Sequence 975, App
29	105	1.4	258	1	US-10-512-184-26	Sequence 26, Appl
30	105	1.4	327	1	US-10-512-184-62	Sequence 62, Appl
31	105	1.4	327	1	US-10-512-184-64	Sequence 64, Appl
32	105	1.4	589	1	US-10-512-184-63	Sequence 63, Appl
33	105	1.4	589	1	US-10-821-234-1494	Sequence 1494, Ap
34	105	1.4	868	1	US-10-821-234-1082	Sequence 1082, Ap
35	104.5	1.4	251	7	US-11-054-515-1810	Sequence 1810, Ap
36	103.5	1.4	509	1	US-10-821-234-1093	Sequence 1093, Ap
37	103.5	1.4	514	1	US-10-821-234-998	Sequence 998, App
38	103.5	1.4	604	1	US-10-793-626-390	Sequence 390, App
39	103	1.4	261	1	US-10-512-184-35	Sequence 35, Appl
40	103	1.4	643	1	US-10-510-386-8	Sequence 8, Appl
41	102.5	1.3	252	1	US-10-512-184-28	Sequence 28, Appl
42	102.5	1.3	1107	1	US-10-485-517-145	Sequence 145, App
43	102	1.3	1970	1	US-10-821-234-1641	Sequence 1641, Ap
44	101.5	1.3	1187	1	US-10-821-234-955	Sequence 955, App
45	101.5	1.3	1560	7	US-11-059-982-1	Sequence 1, Appl

ALIGNMENTS

```
RESULT 1
US-10-821-234-1120
; Sequence 1120, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Steche-Crain, Birgit
; APPLICANT: Andarmati, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: Pf. Seq_Genes Version 1.0
; SEQ ID NO 1120
; LENGTH: 4384
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1120

Query Match 2.5%; Score 191; DB 1; Length 4384;
Best Local Similarity 18.1%; Pred. No. 0.00016;
Matches 303; Conservative 242; Mismatches 613; Indels 518; Gaps 76;

69 SLEAPELLIKRLGSDFNATATVLPVHVIDRLGDSRDTVREKAQLLRDLMEHRVLPQNL 128
913 SLGASASLRPSSD-RSYT-----LNRSYARDSMIE-ELVPSKEQLLTFTRRD 963
129 IDKL-----ATSCFKKN--AKVREEFLQI-IVNAL--HEVYQQLSVRYIYP----- 172
964 SDSLRLHYSWADTLDDNVLVSPHSGFLVFMVDARGSMGSRHHGRIILIPRCKTA 1023
173 ----VCALL-----GDTVANREALITQVLEIYHGVDRLRPDLRRMDVV----- 214
1024 PTRITCRLVKHKLANPPEHERRGISRLVEMGPAGQFLGPVI-----VEIHFQSMRCK 1080
215 ASKLAML-----EOKPDVQKOEGLLPSALKNTNGNGVGLDEADNIGIRERTMI 265
1081 ERELVLSENGETKKEQFDS-KREDL-----TELANGMBEIDSPBELG-KKRICRII 1133
266 KR--PLHSAVSSSLRPKN-VNDVTGAGAVTWSSFESEF-----EYVQ 307
1134 TKDPQYFAVVSRIKQESNQIGPEGILSSTTVPLVQSPFEGALTKRIRVGLAQAPV- 1192
308 LNFPAKMDIYKQVL-----VIISDKMDEMGRV----- 338
```



```

Db 1193 -----DEIVKILGNKATFSPVIVTEPRRRKHKPTWTITVPVPPSGEGVNGYKG 1243
Qy 339 DALKIRALLISYHTPOFVANOLKELS-----LSFV-----DITL 374
Db 1244 DTTNRLILCSITGGSP-----AQMEDITGTTPLTIKOCVSTTNVARSFPLADCHOVL 1299
Qy 375 KE-BERSQVIRACITIAVMSK-----TLRNKLDAPCWS--ILEHLINILIONSAYV 422
Db 1300 ETVGLATQLYRE-LICVPYMAKVVFAKNNDPVESLRCFCMTDDKVDKTLBOQENFEVB 1358
Qy 423 IASASTIALK---YIKYTHAPKLIK-----YTTDLAQ---SKSKIRSTLCCL 466
Db 1359 AASKDIEVLEGKPIYVDCGNLAPLTKGQOLVFNFYSFKENRLPFSIKIRPTSGPCPR 1418
Qy 467 MWLFEEMOTKALERNATV-LRDTL-----KKSIGADCCARRHSR---YAWAPFRHPE 518
Db 1419 LSPLEPKTKKLPOTAVCNLNTLPAKKETESDODDEIEKTRROSFALAKRRYSY 1478
Qy 519 LADQIYGLDIAAORALEREREGGGGTGTGTAPETRTVSRIGRTPTLOKPTPSM 578
Db 1479 LFE-----PGWIERSTGAT 1492
Qy 579 RSISAVDTAAOARAKVRAOYLYSRORKPLGNNSNOA-----SMTGAASGSILPRRL 632
Db 1493 RSLP-----TYYS--YKPFSTRPYOSWTTAPITVPGAKSGF--TSL 1531
Qy 633 NSNSGTPATTP-----GSVTPRP-----RGRAGVSOQSGSTSPSTLAROYGI--- 680
Db 1532 SSSSNTFASBLKIMSVSTSPFKSLTGASTTSVKSISDVAFISLRMTSPKIV 1591
Qy 681 -----GNYNR--GATGALPKKASGIPRSTASRETSPTSGGGLMKRMSYSG 726
Db 1592 VQSPYNIQVSSGTLARAPAVTEAPPLK--GLASNSTFSRSTPVTGASLLERSIT-- 1647
Qy 727 ACSRTTPERNNVRSAPARLLAOSREAHNTLGVDDGCPDYVSGDYMRSGMGRKLM 786
Db 1648 ---WTF---PASPKSNINMSSSLPFK-----SIITSAPLLSSPLK 1683
Qy 787 G-----RDESDIDEASVCSERSFDSYTRGNKNSVLSGSHRDLMDSTORAPFDIE 841
Db 1684 SVSVKRVADVTSAKITMASLSLSPVKOMPGHAEVALVNGSIFPLKASS-----S 1736
Qy 842 TIIQCASTHMSERKQGLISLTQYADGKELTQOOLKCVLDMFRKMDTHTKVSLFLD 901
Db 1737 TLINCKCAT-----ATLQEKISSATNSVSVSAATDVEKVFSTTAMPSPSIRS 1787
Qy 902 TVTTELILVHANETSRRNGSSCLTRLENKLGTDLNSHMSKTKTQVYHEVFTQLOKE 961
Db 1788 YVS--AAPSASFQSLRTPSASA---LYTSLGSS--ISATTSSVTSIITVPVSVVAVLPEP 1841
Qy 962 LFRISDSTQPTTTRTALILRFLDLA-----NRYCCSSDRPSQOQACERTVUKLOL 1016
Db 1842 ALKKLPDONS--FTKSAALLSPIKTLTTEHPHPFSTSPVKSJFLABSLKLTSP 1899
Qy 1017 AADQKMEILSQARSC-----LVALLYNLTPQMTLLADLP-----XYV 1055
Db 1900 SSLSSSQGLKOVAMKEDLMWTALILTQDVEKRPQELPKERIDEBEFKIVENK 1959
Qy 1056 QDSAR-----SCIHSMRRQOSCNSGANSPSSPLSSSPKPLSPVGVPAASLOS 1107
Db 1960 EDLVVSEILKKDVCVNDKSPKSPKSDG--HSPEDDDIEFSESEIREA----- 2007
Qy 1108 HHHOLISSTSPRSQSSVEOELLPSSELDIOHNIOKTSEIRHFGQGYOTALANGFN 1167
Db 2008 --RQDAASQSP-SLPERYQVAKAASEKD--YNLTKYIDVTNIGSSSLT----- 2054
Qy 1168 GHLQY-HDQGOQDSCASLSSNKTOS SANTTOSNTPEASATMRDLNLEB----- 1215
Db 2055 -NLKTKFEDAKKDGQKRVLKPAIALQEHKLMP--PASWRTSTSEKELCMADSFPGT 2112
Qy 1216 -----RTTQAKSPTDAKIVYS--INMAENGELILASNLMESEVVRALTLT 1262

```

```

Db 2113 DTLESPPDSQHDQKSPSLSDSGFETRSKTPSAPQSAETTGPKPLFHEVPIPVITET 2172
Qy 1263 KDQVYELLQTSILTNIGLICIKGNCGLPN-----KHFRS 1295
Db 2173 RTEVHVHVRSDYPAG-----DVPOQDEBPVSPKSPPTFMLEBPKPTTSSIKENK 2224
Qy 1296 IMRMLNLLEAHTDVAIAGLHVLSTKIMRSKMRHNNHFLLELLIKTIQCYOHSKALR 1355
Db 2225 AFQMKASSEEDDH-----RVLSKGMRVKEETH-----ITTTBRVYHSPPGE 2268
Qy 1356 DIDSMPILASLPUDISINIVNPIVIANGEFPTLCAKILLEVTEHHGSEITDAHLIV 1415
Db 2269 GASERIEE-----TMSVHDIMKAFOSGRDS-----KELAGJFEHKA-----V 2307
Qy 1416 FPNLARSADDTQSMWRKAACVIAVLYVLEGEKVPPLSYLNPSTKVLNLYIEK 1471
Db 2308 SPDVHKSALETISQAHER-----DNQMKPLLE-----RIIEVHIEK 2343

RESULT 2
US-10-485-517-252
; Sequence 252, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; TITLE OF INVENTION: Antigenic polypeptides
; FILE REFERENCE: P10062960
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 252
; LENGTH: 761
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-485-517-252

Query Match 2.5%; Score 186.5; DB 1; Length 761;
Best Local Similarity 18.3%; Pred. No. 2.5e-05;
Matches 136; Conservative 129; Mismatches 358; Indels 119; Gaps 19;

Qy 545 GGTGTGTGTAETRTVSRIGRTPTLOKPTPSMRSISAVDTAAQRAKVAQYTLVSRQ 604
Db 81 GSTSTLSNSTRSGSTIS-----TSTISBSTSTFKSSVSTSLMSSTSTSLDSTSLSTS 136
Qy 605 RKPLGPNNSNOASNTGAAAGSLPRPLNLS-----GCTATPTGATTPPRGRA----- 655
Db 137 LSDSTSDSKSGLSTMSNSTSDSTSTSKSDSTSTSTSLSGSTSESDSTSSSEKSDSTS 196
Qy 656 ---GVSGOPCSRSTSPSTKLKQDYGGIGNYRGATGALPKKASGIPRSTASRSTSPTR 712
Db 197 MSIMSOSTSGSTSTSTSTSLSDS-----TSTSLSLASANO 233
Qy 713 SGGGLMRSMYSTGAGRRTPERNNVPSPAPARLLAOSREAHNTLGVDDGQDPDYVSGD 772
Db 234 SGVDSNAGSAGASNSTSTSTSESDQSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 290
Qy 773 YMRSGMRMRKMLMGRDESDIDSEASVCSERSFDSYTRGNKNSYLSG----- 823
Db 291 TSGGSAVSTASLSGSESDSOSISTSAESTSEASSTSLSDSTSTSTSTSTSTSTSTSL 350
Qy 824 -----SHTRLDMST-----QAPFDIET--IIOCASTHMSERKQGLISLTQYLA 867
Db 351 NSASASDSLSSTSLSDSTSTASMSQSSSDSOSTSASLSDSLSTSTSNRMSTIASLSTSVS 410
Qy 868 DGKELTQOOLKCVLDMFRKPMFMDTHTKVSLFLDTVTTELILVHANETSRRNGSSCLTRLF 927

```

```

Db      411  TSESGSTSESTSESDSTSTSLSDSGST-----SKSTASGSASTSTST 454
Qy      928  NKLCTDLNNSMNSKI-----WKTLQVHNEYPTOLQKELFRIISDSTOPT--TKTRIA 980
Db      455  DSRSTSASTSMRSTSDSGSMSTSTSTSMSTSTSLSDSVSDSTSDSTSASTSGSMS 514
Qy      981  ILRELTLANTYCKSDPPS---DQSOACERTVLMIAOLAADQKSMELRQARGLVALY 1037
Db      515  VSISLSDSTSTSTASAEVMSASISDQSMSESVDNSESVSNSBDEKSMGSGTSGV- 572
Qy      1038  NLNTPQWTLMLADLPKYQDSARCTSHMRROGSCNCSGANGSPSSPLS-----SSSPK 10922
Db      573  DSGSLSVSTSRKSESVSESSSLC-----SQSMDSVSTSDSSLSLSTSLSRSES 624
Qy      1093  PLOSPSVCPASLO---SHHHQLSSTSPSPRSQSVQELLFSGELDIQNIQKT-SEE 11448
Db      625  VSESDSLSDSKSTSGSTSTSTSGSLSTSTSLSGSESVESLSTSLSDSJMDSSTSDSDS 684
Qy      1149  IRRHFGGQYOTALAPNGFNGLHQLYHDQOQDSCALSSNSKTOSGANTQNTPEASATMR 1208
Db      685  LSGSLSLSGSTSLSTSD-----SLSDSKSLSSQSMGSGSESTSTSVSDSQSS--TSMNQ 737
Qy      1209  LDNLERERTONAKSPTDPAKY 1230
Db      738  FDSMSISASBEDSMSTSDSSST 759

```

```

1      RESULT 3
2      US-10-485-517-212
3      Sequence 212, Application US/10485517
4      Publication No. US20050256299A1
5      GENERAL INFORMATION:
6      APPLICANT: University of Sheffield
7      APPLICANT: Biosynexus Incorporated
8      APPLICANT: Foster, Simon
9      APPLICANT: Mond, James
10     TITLE OF INVENTION: Antigenic Polypeptides
11     FILE REFERENCE: P100629WO
12     CURRENT APPLICATION NUMBER: US/10/485,517
13     CURRENT FILING DATE: 2004-02-02
14     PRIOR APPLICATION NUMBER: GB 0118825.9
15     PRIOR FILING DATE: 2001-08-02
16     PRIOR APPLICATION NUMBER: GB 0200349.9
17     PRIOR FILING DATE: 2002-01-09
18     NUMBER OF SEQ ID NOS: 424
19     SOFTWARE: PatentIn version 3.1
20     SEQ ID NO 212
21     LENGTH: 1448
22     TYPE: PRT
23     ORGANISM: Staphylococcus aureus
24     US-10-485-517-212

```

[illegible]

```

OY 290 AGATWTFSSFEVVEVQLIFPA-----KMDIDYKQVULVISOXNDWERSVD 333
Db 754 TNA-TDREKQOAIKQVDQ-NVQRALESINNGVONGVDALATQ-----GKAIDALOVD 805
OY 340 ALKKIRALLILSYHT-----QPFQVAVQLKELSPFDILKEELRSQVIREACTITTA 391
Db 806 ATVPKPAKNOAIEVKAEDPTEKESIDQSDOLTAEETKEETALAMIKOITDQAK-OGITDA-TTTA 863
OY 392 YMSITLTKNGKDAF-----CSLIEHLINLIQNSAKYIASASTALAKYIIKYTH 439
Db 864 EVEBAKAQGEAEFNIQIDSTEROKAIEETALDQIEAGVNNADATTEBEK---AFIN 920
OY 440 APK--LLKITYDPLNOSKODIRSTCELAVALLFEEMQTKALERNA-TVEBDTKKSI-- 494
Db 921 ALBEDISKATEDISDQTTNMAEIAVTKNSALEQKAGINEPVKKNALEAREVNNQKIEI 980
OY 495 ---GDADCDARRSRVAYMAFRHFPELADQIYGT-----LDIAQRLERE 540
Db 981 IKNADADASAKEIAR--TDLGRYFDFEADKLKTQTTNAEVALQNTVIAIEAIVBQND 1037
OY 541 -----GGGGGGTGCTGAPETRRTVSIRG-----TPGTLQKPTPS 577
Db 1038 PDANDTNGIDNDATANSNANAATPENTG-QPVVSETTAN-GRADASPTTPNNSDAATGE 1095
OY 578 MRSIAVDTAQAQAKVRAQYTLYSROKPLEGPNNNSQASMTGAAASGS--LPRPRLNSN 635
Db 1096 TTATSAITDDA-----NDKQOANNSSVDASTNSPTMNDVNTSKAPEVST 1139
OY 636 SGGT--PATTPGSVTPRPRAGVSVOSQPGSRNSTPSTYLRDQYGGIGNYRGATCAIP 692
Db 1140 NNGTIDKPVLETIDNATP-----AESTNNNSTTATYN-----ENAPGTSTATAP 1183
OY 693 KKAAGCIRSTASARETETPTSSGGIMKMSVSTCAGSRRT-----PERNNPVY----- 740
Db 1184 TTASTTEASGASDSXDNASVDS--KONAEVNNASBESQSTNDKYAQPKSEKKAEXDGS 1240
OY 741 PSAFAIRLAAOSREAEHTLGVDQOPPYVSGDYMRSGMGEMGRKLMGRDESDDDIDSEASS 800
Db 1241 DSTQNSVESTETLEPADIITEPVNPENTKX-----KEESTNQDTAQO 1285
OY 801 VCSERSFDSSTYRGNKNSYLSGSHTRPLDWSFORAPPDIDETIIQFCASTHWSRKGGLI 860
Db 1286 LKSETTNVAS-----NEADKSPSKADTEVSNKPTSSASEAK---EKMTSTNVASOKDPTAT 1337
OY 861 SLTQYLDAGKELTQOOLKCYLDMRKMFMDTHTVYSLFDVTYTELILVHAN-ETSRNGS 919
Db 1338 ADTN-----DTQKSVGSAANNKATQNDGANCASPATVSNGS 1372
OY 920 SSCLTRLENKLGTD 933
Db 1373 NSANQDMNTANTNTD 1386

```

```

/ RESULT 4
/ US-10-971-982-2
/ Sequence 2, Application US/10971982
/ Publication No. US20050244889A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosenfeld, Michael G.
/ APPLICANT: Glass, Christopher K.
/ APPLICANT: Rose, David W.
/ APPLICANT: Torchia, Joseph
/ TITLE OF INVENTION: A Transcription Factor Coactivator Protein, P/CIP
/ FILE REFERENCE: 6627-P/Al021
/ CURRENT APPLICATION NUMBER: US/10/971,982
/ CURRENT FILING DATE: 2004-10-21
/ PRIOR APPLICATION NUMBER: US/09/445,353
/ PRIOR FILING DATE: 2000-05-05
/ PRIOR APPLICATION NUMBER: PCT/US98/12263
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/049,452
/ PRIOR FILING DATE: 1997-06-12
/ NUMBER OF SEQ ID NOS: 14

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1402
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3121)..(3121)
; OTHER INFORMATION: "n" is any nucleotide
US-10-971-982-2

```

```

Query Match      2.0%; Score 153.5; DB 1; Length 1402;
Best Local Similarity 18.6%; Pred. No. 0.0095;
Matches 276; Conservative 196; Mismatches 557; Indels 451; Gaps 69;

```

```

QY 126 QALIDKATSCFKH-----NAKREFEFLQTVNALHEGYTOUS-----V 166
DB 6 ESSLDPLAESKRLPCAPGGLVYSGEKMRROESKYIELAELISANLSDIDNFV 65
QY 167 RYIIPVCALLGDPVNVNREAAIQTLVEIKYKVGDRLEBRLRMDVDPASKLAMEQKD 226
DB 66 K---EDKAILKETVRQIRQ-----IKEGKTISD---DDVQKADVSTGCGVI 109
QY 227 QYKOGGILLPSALK-----NTNGNGVGLDEADNIGLRERPTRMKIKPLHSAVSSSLRP 280
DB 110 DKDSLGLPLLQALDGFLEFVNVRDGNIVFSENVTOYLOKQEDLVNTSVSYSLHEPRR-- 167
QY 281 PNVDVDTGAGAVTESFESEFVVPOLNIFPAKMDMDIYKQVLIIISDKNADMEKRYDA 340
DB 168 -----KQFLNTQY-NPOLMEFLG-----LWRTDKKAPYLIIYRM 201
QY 341 LKTRALLILSYHTPOFVAVQKELSLSEVDILKE---ELRSQVIREACTITAYWSK 395
DB 202 LMKTH-----DILEDVNASPETRQRYETMQCFALISOPRA 235
QY 396 TLRLNLDAFCWSILHLINLIONSAKVIASASTIAIKYIKTHAPKLIKITYDTLNSK 455
DB 236 MLEBEGEDUQC-----MICVARRVTPAPPSPEPSFTIRHDLGKVVNIDTNSLRSSM 287
QY 456 SKDIRSTLCMLVLLFEEMQTKALERNAVLRLTLKK--SIGDADCDARRHSRYAVMAFR 513
DB 288 RPG-----FED-----IIRCRQIRFSLNDGGS-----WSQK 314
QY 514 RHPEPLADQIYGTLDIAQALEREEREGGGGTGTGTABETRTVSRIGRTPT--- 570
DB 315 RHYQGA--YVHGHAETPVYRF-----SLADGTIVSAQTK---SKLFRNVTNDR 358
QY 571 -----LQK-----PTPSMSISAVDTAAQRAKAVRAQYLLYSRQRP----- 607
DB 359 HGFISTHFLQREONGYRNPPIPODKGI-----RPPAGCGVS 395
QY 608 LCPNNSNOASMTGAASGSLPRRLNSNGGTPATTPGVTPRPRGAGVGSOSQGSRT 667
DB 396 MSPNONGV--MMGSTTYG--VPDPSTGQMGGA-----RYGASSVASLTGQSLQ 442
QY 668 SPSTKLRLQYQ--GIGNYRGATGAI PKASGI PRSTASRETSPRSQGLMKRSMYSTG 726
DB 443 SPSSVONSYSGLSMSSPHGSPGLAPNQONIM-----ISPRNGSPKXASHQFPA 493
QY 727 AGSRTPERRANNVYRSAPAKLLAQSREAEHL-----GYGDD----- 763
DB 494 AGA-----HSPMGSGNTGSHSFSSSSLSALQAI SEGGTSLTSLSPGPKLDNSPNM 547
QY 764 --GQDPVYSGDYMRSG--GWRMGKRLMGRDESDIDISEASSVCSERSPDSYTRGNKSNYS 820
DB 548 NISQPSKYSGGDSKSPGLIYC-----EQNPVE---SSVQSNRSPQVYKESKES-- 594
QY 821 LSGSHRLDMWSTORAPFDD--LETTIQF--CAST---HWS-----ERXDLGISLTQY 865
DB 595 -SGEVY---ETPRGPLESKGHKLLQLTLTSSDDRGSHSLTNSPLDPKCKOSSVAVYS- 648
QY 866 LADGKELTQOQOKCVLDMRKMFMTHTKVVYSLF--DVTVELLVHANENSRNCS--S 921

```

```

DB 649 PSGVSSSTGTVSSTSNVHSGLQEKRIILKLLQNGNSPAEVAKITAEATGKDTSTAS 708
QY 922 CLTRLFNKLGTDLNLSMHSKIMKTLQVNHXYF-----PTQLQKEL----- 962
DB 709 C-----GEGTTRQEQSLPK--KKNENNALRYLLDRDPSDVIAKELAQADSGSKLSQC 761
QY 963 ---FRISDSYQTPPTTKRI-----ALRFLTDLANCYCSSDPSDSQACE 1007
DB 762 SCSTNPSSGQEKDKRIKTENDEVSGDLNDLALIGDLT-----SSDFYNNFTNGCH 813
QY 1008 RTVLKLAQLAADQKSMELRS--QARSCVALYNN-----LNTPQMTLLADLPKYQDSA--- 1059
DB 814 PGAKQ--QMFGRPSLSGLRSFPQVQSVPRPNRAVLSDSP---YSVSGPVPKXVSAFPG 868
QY 1060 --RSCISHMRKROSQSCNSGAN-----SPSSSP----- 1085
DB 869 LPKQPILAGNPRMMDQENYGANMGNPRNVNPTSSPGDWGLANSRASRMEPLASSPLG 928
QY 1086 -----LSSSPKPLQSPSVGPFASLOSHHQLSISSTSPRSROSSVEQELLFSSBLDIOH 1140
DB 929 RTGADYSATLPRPAMGGSV--PTLPLRSNR-----LPGARFSLQOQOQOQOQOQOQ 979
QY 1141 NIOKTSERIRRCFGQOYOTALAPNGFNGLQYHDQGOQDSCASLSNSKXTOSANT----- 1196
DB 980 QOQOQOQOQMLQMRGELPRMGVAVPYSPAVSQNGPSMPE--GMSMEGPHGSOVRPILR 1038
QY 1197 -----TQNTPEASATMR--LDNLER--ERTONAKSPDDAKVITVSINMAENGELI 1244
DB 1039 NSLDLGLGPFENAGQDERALLDQLHTFLSNTDATGLEGIEDRALGIPELVNOQO--ALE 1096
QY 1245 LANSMESEVVRVALTILTKQPVVELLOTSLNLGICIKG--NCELPNHFPSIMMLNI 1303
DB 1097 SKQDVFQGEAAVNM---DQKALYQOTYPAQGPPLQGGFNLDGQSPFSNMGOI--- 1149
QY 1304 LEAHTDVYAGLAVLSKIMSNKVRHMMHFLLEILIKIQCYSKEALRDISMIPR 1363
DB 1150 --SQGSPFLQGMTPRAGVPR---RINTPKQLAMOLOQRLOQGOFLNQSQALEMKEN 1204
QY 1364 IAPSLPLDLSINIVPIANGEPFTNLCAKILLLEVTEHH 1403
DB 1205 PAGTA-----VVRPMMPQAFNAQMAA--QCKELMSHH 1236

RESULT 5
US-11-055-035-1
; Sequence 1, Application US/11055035
; Publication No. US20050256072A1
; GENERAL INFORMATION:
; APPLICANT: AKONIN, NEIL
; APPLICANT: ZAMORE, PHILLIP D.
; APPLICANT: BRODERICK, JENNIFER
; TITLE OF INVENTION: DUAL FUNCTIONAL OLIGONUCLEOTIDES FOR USE IN REPRESSING
; FILE REFERENCE: IMY-095
; CURRENT APPLICATION NUMBER: US/11/055, 035
; PRIOR FILING DATE: 2005-02-09
; PRIOR APPLICATION NUMBER: 60/543,467
; PRIOR FILING DATE: 2004-02-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 1
; LENGTH: 3144
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-055-035-1

Query Match      1.9%; Score 146.5; DB 7; Length 3144;
Best Local Similarity 17.7%; Pred. No. 0.089;
Matches 289; Conservative 213; Mismatches 527; Indels 605; Gaps 71;

```

```

QY 193 VEIKYHVGDRLRPDLRMDVPAKSL--AMDE-----OKFQVQKQGLLPSALPNT 242
DB 131 MELFLTCSDAESVYKVADECAKLVIKALMDSNLPRLOLELYEIKKNG--AARSLR-- 186

```

QY 243 NNGVGLDEADNIGLERPTMIKRP-----LHSAVSSSLRPKNVNDVTGAGAVTMS 297
 DB 187 ---AALMFPAELALVBPQKC--RPYLVNLLPCLTRTSKRPDESQETLAAVPKIMAS 240
 QY 298 PESSFEVVPQNLIFAKMDDDIYKQVLVIISDKNMDMEKRVDAKKI---RALLLSYHT 354
 DB 241 FGN-----FAND-----NEIKVLKAFIAALKSSPTIRTAAGSAVSIQOHR 284
 QY 355 QGFVAVOLKELESLFVDILKEELRSQVIREACITIAVMSKTLRNKLDFAFCSILEHLIN 414
 DB 285 RQYVYSMLNVLGLLVEDEBHTLLGLVLTIRY-----LVP 325
 QY 415 LIQNSAK--VIASASTIALKYIIKYTHAPKLIKITYDTLNQSKSDIRSTICEMLVLLFE 472
 DB 326 LIQQVQKDTSLKGSFCVTRKEMEVSPSAEQVLQVVELTLHHTOHQ-----H 372
 QY 473 EMQTALENNATVLRDTLKKSIGDADCDARRSRVAYNAFRHFPRLAQIYGTIDIAAQ 532
 DB 373 NVVTGALF---LLOQ-----LFRTPPPELLQTLTAVAGIGQL 406
 QY 533 RALERESEGGGGGT-----GNGTGTAP-----ETRTVSR 564
 DB 407 TPAKESEGGRRSGSIVELIAGGSGSPVLSRKQKXVLGBEBALDDSESRSVSS 466
 QY 565 GRTPGTLQKPTPSMRSISAVDTAAQRAKV---RAQYTLYSRQKPLGPNNSNOAS-- 617
 DB 467 ALTAIVKQIBISGBLAASSGVSTPGSAGHDIIITEQPRSGHTLOA-----DSVLDASD 518
 QY 618 MTGAASG-----SLPR--PRLNSNGGTPTATPGSVYTPRRGAGVQSOP 662
 DB 519 LTSSATDGEDDILSHSSQVSAVPSDPAMDND--GTOASSPSISOSQ-----TTTEG 570
 QY 663 GSRSTSPSTKLDQXGIGNYRG-----ATGAIPKASGIRSTA----- 703
 DB 571 PDSAVTPDSSEIYLDGTNOYLGLQIQOPQDEDEATGILDEASEARNSMALQQA 630
 QY 704 -----SSRETSPTRSSGG-----LM 718
 DB 631 LKKNMHCQPSDSSVDKFLVRLDEATEPDQENKFCRIKIDIGQSTDDASAPLVHCVRLL 690
 QY 719 KSMVSTGSGSRRTPERNNPVPSPAPALLA-----Q 750
 DB 691 SASFLITGKKNLVLPDRD--VRVSVKALALSCVGAVALHPESFSLKYVPLDTTYPE 748
 QY 751 SREAEHTLGVDDGQPD-----YVSGDYMSGGMR----- 780
 DB 749 EGYVSDILNYIDHGDQVNGATAIILGTLICSLISRSRHVGDMM--GITRILTNTFSL 806
 QY 781 ---MGRKLMGRDESD-----DIDSEASSVCSESRFDSYTRGNKSNYSLSGSHR 827
 DB 807 ADCIPILKRTL--KDESSVYCKLACTAVRNCVMSLCS-----SSYSELGLQI 852
 QY 828 LDMSTORAPFDIETIIQFCASTHSEKRD----- 857
 DB 853 IYVLTIR-----NSSYMLVTRTELETTAEIDFRVLSFLAENLRGAHHY 899
 QY 858 -GLISLTQYIAD-----GKE-----LTOQQLKCVLDMFKMDHTTKYVSLFLD- 901
 DB 900 TGLLAKQERVLNNVILHLGDEDPVRVHVAASLILVPLKLFYKCDQOGADPVVAARQD 959
 QY 902 -TVTELLIVANETSNSSSCLTRLFNKLGTDLNS-----MHSKIMKTLQV--HEVP 954
 DB 960 SSVYKLMLHETOPRPHFSVSTIRYR--GYNLLPSIDVTMENNLSRIYAVSHELIT 1017
 QY 955 TOLQ-----LKEIFRII-----SDSTQPTTKTRAILRFT 986
 DB 1018 STRPALTFCCCEALCLSTAFVPCINSLGHCVPPLSASDESKSCVGMATMILLIS 1077
 QY 987 DLANTYCKSSDFSDSOACERTVVLKLAQADQKSMELRSQASCLVLYLNTPOMTL 1046
 DB 1078 ---SAWFPLDLS--AHQDALILAGNLLAASAPKSLNSVASEBEA-----NPAATK 1123

QY 1047 -----LLAD---LPKYQDSARSCISHMRQOSQ-----NSGANS 1080
 DB 1124 QEEVWPAIGDALVPMVEQ-----LPSHLKVINICAHVLDVAPGPAIKALPLSTNP 1177
 QY 1081 PSSSPL-----SSSSPKPLQSPSVGPASLQSHHQLSISSTSPRSQSV- 1127
 DB 1178 PLSPIRKKGKEKEGEGASVPL--SPKGSASMAAS--RQSDTSGPVTTSSKSSLSGSFYH 1234
 QY 1128 -----OELL-----FSSBLDIQANIQKTSBEIRHCFGQOYQALAPNGNHLQYHD 1174
 DB 1235 LPSYLKLDVAKATHANKVTLIDJNSTEK-----FGGFLREAL----- 1273
 QY 1175 QGQODSCASLSNSKGTSSANTTQ-----SNTPEASTMRDNLERERTONAKSP 1224
 DB 1274 ---DVLSQIIELTALQDQICVEBILGYLKSCTSRPMATVCCVQQLKTLFGNLSQ 1329
 QY 1225 TDDAKVITVSINMANGELILASINMESEVVRVALTLTKDQVPELLQSLTNLGI---CI 1281
 DB 1330 FDG-----LSNPSKQ-----GRAQRLGSSSVRPGLYHYCF 1361
 QY 1282 KQNGCELPKPRSIM--RMLNLTLEAHTDVIAGLHVLKTI---MRSN-----KMR-- 1329
 DB 1362 MA-----PYTHPTQALDASLRNMVQAEQENDTSGMPVLQKVSQTLKTNLTSTKNRAD 1416
 QY 1330 ---HNMHPLELILKTIQ-----CYHSEKALRDIDSMIPRIAPSLPLDSINIYVP 1379
 DB 1417 KNAIHNRILPEPLVTKLKQYTTTTCVQLOKQVLDLQVLQVLRVNTCLDSDQVFIG 1476
 QY 1380 V-----IATGEFP-----TNLCAIKILEVTEHGESEITDAHLDIYFNPILARSADTQ 1427
 DB 1477 VLKQREYIEVQGFRESEAIINIFPVLVLSYERHSQI-----IGPKIILQCDGIM 1530
 QY 1428 SMVRKAAVFCIVKL 1441
 DB 1531 ASGRKAVTHAIPAL 1544

RESULT 6

US-10-793-626-2964
 ; Sequence 2964, Application US/10793626
 ; Publication No. US20050255478A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KIMBERLY, WILLIAM JOHN
 ; TITLE OF INVENTION: STRAPHLOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
 ; FILE REFERENCE: PUS4800S
 ; CURRENT APPLICATION NUMBER: US/10/793,626
 ; CURRENT FILING DATE: 2004-03-04
 ; PRIOR APPLICATION NUMBER: 60/164,258
 ; PRIOR FILING DATE: 1999-11-09
 ; NUMBER OF SEQ ID NOS: 4472
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2964
 ; LENGTH: 5024
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: synthetic
 ; OTHER INFORMATION: amino acid sequence
 ; FEATURE:
 ; NAME/KEY: MOD RES
 ; LOCATION: (5024)
 ; OTHER INFORMATION: variable amino acid
 US-10-793-626-2964

Query Match

1.9%; Score 146; DB 1; Length 5024;

Best Local Similarity 17.9%; Pred. No. 0.19; Matches 279; Conservative 233; Mismatches 590; Indels 456; Gaps 66;

QY 88 TATVPLVHYDR--LQSDRDYREKALQILRLDMHRVLPALIDKATSCFKAKNAV 144
 DB 2876 TRSEVHOYINRQALNDSMTLRQS-----ITD--EHEVQTSNYINETVGNQTAYNNAVD 2929
 QY 145 REEFLQIVNALHEGTQOLSVRVYIIPVICALGDPYTVVRAEAQTLVEIKHVGDRLR 204

```

Db      | 2930 R---VKQIINQ-----TSNPTNMPLE--VERATSNVKTSKDALKH 2963
Qy      | 205 PDLRRMDVBPASLAME--QKFDQVKQEGC-----LLPSALKRTNGVGLD----- 250
Db      | 2964 GE-RELNDKXSKTFVAVHLDNLNOAQKALTHEIQATIVSOVNNIYAKAKALNNDMK 3022
Qy      | 251 -----EADNIGLREPRPTMKRPLHSVAVSSSLRPPNVNDVTGAGAVTMSFSSFEV 304
Db      | 3023 LKDIYAQOQNV-----ROSNNYINEDSTPQNMVND-----TIMHAQSIIDQ 3063
Qy      | 305 VQOLNIFAKMDDDIYKQVLVITSDKA--DMEKRVNALKKIPALLILSH--TORQFVAV 361
Db      | 3064 VANPTMSH---DEIENAINNNIKHAINALDGEHKQOAKENANLLINSLNDINADPORDAI 3119
Qy      | 362 QKKEISLSEFVDILKEELRS--QVIREACITAVMSKTLRNKLDFAFWSILEHILNIONSA 420
Db      | 3120 NBLVNEAQREKVAEBOLSAQALNDA-----MKHLRNSIQONS 3157
Qy      | 421 KVIASASTALKYI-----IKYTHA-----PKLLKIYDTLQNSKS- 456
Db      | 3158 SVRQES-----KYINASDAKKEQYNHAREVENIINEQHPTLDKEIHKQLTDAVNOAND 3212
Qy      | 457 -----KDISTCEIMVLFEEWOTKALER--NATVLRPTLKSIGDADCDAR 503
Db      | 3213 LNVGELLDADKONAHOSIPTLMHL--NOAQONALNEKINNATVRAKVAALIQAK----- 3265
Qy      | 504 HSRVAYMAFRHFPBLADQIYGTLDIAQALEREGERGGGGGTGTGTAPETRRT--- 560
Db      | 3266 -----ILDHAMENLEESIKDKQOVQOSSNYINEDVDQETVYN 3303
Qy      | 561 -VSIRIGRTPGTLOKPTPMSRSIS-AVDTPAAQAKVRAQYTLY-----SRQKRLPGPN 612
Db      | 3304 AVDHTEILNQTIVNPTLSIEDIEHAINEVQAKKOLRGKQKYOTGIDLADKELSKLDLT 3363
Qy      | 613 SNOQMTGAAAGSLPRPLNSNSGCTPRTTGSVTPRRGAGVQSQSPGSRSPSK 672
Db      | 3364 SOOSSSI-----SNOIYTAKT-----RTEVAQALEKKSINHMWK 3398
Qy      | 673 LRDYGGIGNYRGATGAIIPKASGIPRSTASSRETSPTSGGLMKSMSYSTGASRRT 732
Db      | 3399 -----ALNKIYKADKVLDSRINEDQPEKA---YQQAINHVDIIR- 3440
Qy      | 733 PERNNVPPSAPARLLAQSREAEHTLVGDDQDPYVSGDYNSGGMGRKRL--MGDE 790
Db      | 3441 -QTNEMPTVINSITHELETAQNNL-----HGD-----QCLAHAKQDA 3478
Qy      | 791 SDDISEASVCSER---SFDSSYTRGKXSVYSLSGSHTRLDMDSTORAPFDITITIQ 846
Db      | 3479 ANVINGLJHLNVAQREVMINTNTNATTRKVKAKNDNAQA--LDKA-----METLOQV 3529
Qy      | 847 CASTHMSERKDLISLTQYADGKELTQOQKCVLDMFRKMMDHTKYVSLFLTVL 906
Db      | 3530 VA-----HKNNIILDSKILNEDSKYQOQYDRAVIDAEBOLLQTNPTLEPKYIVIVDN 3583
Qy      | 907 ILVHANETSRRNGSSCLTRLFNKLDGTLNLS---MHSKIMWTLQ-----VHHEYPT 955
Db      | 3584 VL--ANEKILFGA-----EKLSDYKSNANDEIKHMNYLNNAKKOSIKMISHAALRT 3633
Qy      | 956 QLO-----LKEPFIISDSTQTPPTYK-----RIALRFLTDLA 989
Db      | 3634 EVKOLLQOAKTIDEMKSLLEDTOVAVITDTLLPNTYSEDEKKEKVDQVDSHAQAIIDKI 3693
Qy      | 990 NTVCKSPDPSPDO--SOACERTVCLKLAQALADOKSMELRSQARSCLVALYNLTPTMTLL 1048
Db      | 3694 N---GSNVSLDQVQALEQULTQASENLDDGRVBEAKYHANOITDOLTHLNSLO----- 3744
Qy      | 1049 ADLPRVYDASRCSIHSHMRQSQSCNGSANSPPSSSPKPLQSPSVGPFASLOSH 1108
Db      | 3745 -----QOTAKESV-----KNATKLEIATASNMLALNKV--MOKLBOF 3781
Qy      | 1109 HHQOLSISSRPSRSSVBOELLFSSSELDIGNIKO---TSEIHHCQGOVQYTLAENG 1165

```

```

Db      | 3782 INHADISDNRYQADDDKIIAYDADLEHGODIQKSNATONEAKQALQOLINETSJNG 3841
Qy      | 1166 FNGLQYHDQOQDSCASLSNSKTSOSA---NTTOSN----- 1200
Db      | 3842 FE-RINHARPRALEIYISLEKINNAAQSALBDKYQSHDLLEHLVNEGTLNDIMGBL 3900
Qy      | 1201 -----TPESAT---WRLDNLEBERTO---NAKSTFDQAKYITVSINM----- 1237
Db      | 3901 ANAIVNNYAPFKASINYINADNLKKNDFTOAINNARALANTQOQNLDFNAIDPFKDI 3960
Qy      | 1238 -----AENG-ELIASNEMESVVALTLTQDQV-----ELLQTS-----LTLN 1277
Db      | 3961 KTKQALNIEKRLTAKSFAEKLIDSLEKFTNKAQFTHANDEIMNTISIAQLSRIYQADL 4020
Qy      | 1278 GICIGGNCCELPNHFHSIMRMLNLTLEAHTDVAIAGLHVLKIMRSKXKHHMMHLE 1337
Db      | 4021 NDAMKSLRDELNQAFP--VQASSNYINSDE-DLKQQFDHALSN-ARKVLAKENGNLDE 4076
Qy      | 1338 LILKIIOCYOCHSEALRDISMIPRIAPSLPLDSINIYV-----IATGEFTN 1389
Db      | 4077 IOISGLKQVIDETDVALNGIQLRLSKAKAKAIQYVQSLSYINDAQRHIAESNIHNSDDL 4136
Qy      | 1390 LC-----AKIILEVTEHHGSEI-----TDAHLDIYFVPLARSADDTOS 1428
Db      | 4137 LANTLSKASDLDMAMKDLRDTLESNSTSVNSVNYINADKVLQIEFDEALQOASATSS 4194

RESULT 7
US-10-131-826A-10
; Sequence 10, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Guthrie, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria A.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; CURRENT FILING DATE: 2002-04-28
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588

```

PRIOR FILING DATE: 1997-09-19
 Remaining Prior Application data removed - See file wrapper or PALM.
 NUMBER OF SEQ ID NOS: 550

SEQ ID NO 10
 LENGTH: 594

TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-131-826A-10

Query Match 1.7%; Score 133; DB 1; Length 594;
 Best Local Similarity 23.6%; Pred. No. 0.063; Indels 116; Gaps 24;
 Matches 118; Conservative 55; Mismatches 210;

```

QY 443 LKITYTDLN---QSKKDRISTLCELMLL---FE-----EMQTKALBRN 482
DB 133 LYQIYIDLYGGLQKPSDEDEKKLAEKKASIGYTEDSTVAEVEKAAPKEEESAAEE 192
QY 483 ATVLEDTLKKSIG-DADCDARRHRYAYWAFRRHPELADQ--TYGTLD-----528
DB 193 SNSDEDEVIPDIVEDVDLNOEQA-----DLNKQATYGMADGDGVFVRLRKDK 243
QY 529 -----IAORALEREDEGGGGGTGTGTAPETRTVSRIGRTPTGTLQKTPSMRSISA 583
DB 244 EEAERIKIAKALEEKAMYSGRSRRRQRRREFREKLRGKIS-PSYARROSPTIDPYKR 302
QY 584 VDTAAQRAKYAQAQYTLYSRQRKPL-----GNNSNOASMTGAASG-SLPRPLNLSN 636
DB 303 SPSESSSRSRSRSPPTGREKEKIFITISFGSGDEEAAAAAAGSVTTGKKPAPAPQP 362
QY 637 GGTPTATPG-SVTPPRRAGVSGO---PGSRSTSPSTKLADQYGGIGNYRGATGAI 691
DB 363 GG-PA--PGRNASARRRSSSSSSSSASRTSSRSSSSSS--RSRRG--GGYRSGRHAR 416
QY 632 PKKASGIRSTRSSSETPTREGGGLMKMSMTGAGR-----RTERNNPVAPSPA 745
DB 417 SRSRW-SRSRSRSHRYRSRSR-----RGRHSGSGSDGHRYSRPPRRGCGPR---466
QY 746 RLAAQRAEHTLVLGDDQGPYVSGDYKSGMGRKLMGRDESDDIDSEASSVCSER 805
DB 467 ---RSTRSRSH-----SGDKYRRG---GRCLRHHSSSRSSSSSLSPSR 506
QY 806 SPDSSTYTRGNKSNVLSGSHRTLDWSTORAPPDIIETIIIOFCASHWSE-----R 855
DB 507 SLTRSRSHSPSPSGRSRSR---SQSPSPAPAREKLTTPAASPAVGEKLTAKTEPAAGK 563
QY 856 KDGLISLTQYLADEKELTQ 874
DB 564 ETGAQKVTDADSGAEETE 582

```

RESULT 8
 US-10-821-234-972
 Sequence 972, Application US/10821234
 Publication No. US20050255114A1

GENERAL INFORMATION:
 APPLICANT: Labat, Ivan
 APPLICANT: Stache-Crain, Birgit
 APPLICANT: Andarmani, Susan
 APPLICANT: Tang, Y. Tom
 TITLE OF INVENTION: Methods for Diagnosis and Treatment of Presclampsia
 FILE REFERENCE: 821A
 CURRENT APPLICATION NUMBER: US/10/821,234
 PRIOR FILING DATE: 2004-04-07
 PRIOR APPLICATION NUMBER: US 60/462,047
 NUMBER OF SEQ ID NOS: 1704
 SOFTWARE: pc_seq_genes Version 1.0
 SEQ ID NO 972
 LENGTH: 482
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-821-234-972

Query Match 1.7%; Score 131.5; DB 1; Length 482;
 Best Local Similarity 20.1%; Pred. No. 0.059; Indels 255; Gaps 34;
 Matches 127; Conservative 67; Mismatches 182;

```

QY 486 LRDTKKSGDADCDARRHRYAYWAFRRHPELADQYGTLDIAORALEREDEGGGG 545
DB 7 LRD--RRGLDQQTTPARELLSALPHAPR-----QARWPPAASGGGG 51
QY 546 GTGTGTAPETRTVSRIGRTPTGTLQKTPSMRSISAVDTAAQRAKYAQTLYSRQR 605
DB 52 GGGGGWHGF-----GRCP-----PCGDDEDTAAMDSPC-----QP 82
QY 606 KPLPPNNSNOASMTGAASGLPRPLNSNGGTPATTPGVTTPR-RGRAGVSGSGS 664
DB 83 QPL-----SQALPQ-----LPGSSSEPLEPFGARNGVESLYP--116
QY 665 RSTSPSTKLADQYGGIGNYRGATGALPKKA--SGIPRSTASRETPTPSG-----CG 716
DB 117 -----CPLPSTYRCF-----VPSEASAGSCTPRATRTTASPLADFGCGDQGG 161
QY 717 LMKRSMYSTGA-----GSRRTPRRNNPVRESAPARLLA-----QGRE 753
DB 162 EL-RPLQSEGAALVTGCGELAAG--ARPEAPKRYAEDGDGAPSPSKAPWAOENOE 218
QY 754 AEH-----TLQVGD-----DGQPDVYSGY---NRSGRMGRKLMGRDES 791
DB 219 AEREGMNSCSCSSGSEASAGLMEALPSAPERLALDVIYPCMYGICVQDSFLGALG 278
QY 792 DDIDSEASS-----VCSERSFDSSTYTRGNKSNVLSGSHRTLDWSTORAP-----836
DB 279 GRVLAEEVALKRGRLADQGLVSGRAIPRPSRKD-----QIAWEGHEPGGRS 327
QY 837 ---FDIETIIIOFCASHWSEKXGLISLTQYLADEKELTQOOLKV---LDMFRMF 888
DB 328 IGALMAVDAVIRACG-----RLGSYINGR--TKAWACYPGNGLGVRRHV--373
QY 889 MDHTKYVSLPLDTVTLELILVANETSBNSSCLTRFNKLGTDLNL---SNHSLW 943
DB 374 -----DNPHGGRGICITCY-----YLNQNDVVKHGGU--401
QY 944 KTLGVHEVPTQOLKELF-RII---SDSTQ---PPTTKRIALLFLTDLANTYCKS 995
DB 402 -LQIFEGRRVANIEPLFRLLIFMSDRNNPHEVKAIVTRALITWYTD-----451
QY 996 SDPFSQSQACERTVLK-LAQLAADOKSMEL 1025
DB 452 -----AKERAAADKXQLASGQKGVQV 473

```

RESULT 9
 US-10-971-982-3
 Sequence 3, Application US/10971982
 Publication No. US20050244889A1

GENERAL INFORMATION:
 APPLICANT: Rosenfield, Michael G.
 APPLICANT: Glass, Christopher K.
 APPLICANT: Rose, David W.
 APPLICANT: Torchia, Joseph
 TITLE OF INVENTION: A Transcription Factor Coactivator Protein, p/CIP
 FILE REFERENCE: 6627-PAL021
 CURRENT APPLICATION NUMBER: US/10/971,982
 PRIOR FILING DATE: 2004-10-21
 PRIOR APPLICATION NUMBER: US/09/445,353
 PRIOR FILING DATE: 2000-05-05
 PRIOR APPLICATION NUMBER: PCT/US98/12263
 PRIOR FILING DATE: 1998-06-12
 PRIOR APPLICATION NUMBER: 60/049,452
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 3
 LENGTH: 1463
 TYPE: PRT

ORGANISM: Mus musculus
US-10-971-982-3

Query Match 1.6%; Score 125; DB 1; Length 1463;
Best Local Similarity 17.2%; Pred. No. 0.79;
Matches 240; Conservative 194; Mismatches 511; Indels 454; Gaps 56;

```

OY 171 PNYCALLGDPVNVREAAIQTIVEIKYKGDRLRPDRMDVPAKSLAMEQKPDQVQ 230
DB 67 POKCALIKETVKQIRQ-----IKGEKAAANIDEVQSDVSTQGVLDKDA 114
OY 231 EBLILPSPALK-----NTNGNGVGLDEANIGIREPTMIKRP-----HSAVSS 276
DB 115 LGRPMLEALDGFEEFVNVLEGSVFERNVTOYLRYQELMNKSVYSILHVGDFHEFVN 174
OY 277 LARPKNVN--DVTGDAVATMESFESEFVVPQLNIFHAKMDDIYKQVLIIISKNADM 334
DB 175 LIPKMNNGSWSGEPFRSSHTFNCRLVKP-----LPDSEEBG 214
OY 335 EKRVDAKKIRALLISYHTOPFVAVQKELSLSEV---DILKEELRQVIREAC-- 387
DB 215 HDSQEAHQYKEMQCAV--SQPKSIEGEGDLOSCLIVHEDPHBEKTNSSILIRKLYHP 273
OY 388 -----ITAYMSKTLRNKLDACWSILEHLINLIQNSAKVIASASTALK--YIIXY 437
DB 274 GPPRODHFGYHNE-----RHEAGLGRSGKDAFRSSTHMSKGLYHMPR 320
OY 438 THAPRLK-----IYDTLN-----OSKXDIRSLC--ELMVILFEEMQTKALE 480
DB 321 RHHEVELRGGLAFSOTYRPSLSDGTLVAAQTKSLRQSTTNBPQVLISL---HMLHRE 376
OY 481 RNATVLRDITLKSIGDADCDARRHSRYAWAFRRHPELADQIYGLDI---AARAL 535
DB 377 QNVCVANPDLTGQAMGKPLNPISSSSPAHQALCSNPGQDMTLSSNINPMNPKREQGM 436
OY 536 EREERGGGGGTGTGTGTGTAETRRYVSRIGRTPTLQKTPSMRSISAVDTAAQAKXR 595
DB 437 PWGRFGSGG-----MNYSGQATTPQ---G 460
OY 596 AQYTLYSRQRKPLGPNNSQAQMTGAASGLPRPLNSGSGTPTATPGSTP----- 649
DB 461 SNYAL-----KXNSPSQSSPGMNPQASVLSPRQMSGVAGSPRIIPSGSPGNLHS 515
OY 650 -----RPRGAVSOSQPSRSTSPST 671
DB 516 PAVGCSSTGNHSTYNSLSNALQALSEHGVSIGSSLASPLDKMGNLQNSPVNMMPPLS 575
OY 672 KL-----RDQYGGIGNYTRGATGAI---PKKAGIPRSTASSTSETPTRSGGLMKRS 721
DB 576 KKGSLDSKDCFCGLYGEPSKGTGOAEASCHPKKOKG-PVDSMMPQAAAGDRAEG---HS 630
OY 722 MYSTAGSRTE-----RNNVPRPS-APARLLAQRREAHNTLGVDDQGPVDSGDVMS 776
DB 631 RLHDSKGQTKLIQLITTSDDQEPPLPSLSDTNKDS---TSLPGSGTHGTSLE 685
OY 777 GGMWGRKLMGRDESDID---SEASSVCSERSFSSSYTRGNKNSVLSGSHTRLDWSTQ 833
DB 666 KHKILHRL--QDSSSPVDLAKLTAEATKELSSQSSST----- 722
OY 834 RAPFDIETITQFCASTHMSERKDLISLTOYLADQKELTQOOLKCVLDMFRKMFMT-- 891
DB 723 -APGSEV-TVKQEPASPCKKEN---ALRYLLKXDKDQKGLPEITPKLERLDSKTD 775
OY 892 --HTVYSLFLDTVELIVANETSRNGSSCLRLFN--KLGDILNSHMSKIMWKLQ 947
DB 776 ASNTKL--IAKTIVE-----EVSFEPSDQPSGLDLBEITLDLQNSQ----- 817
OY 948 VVHEVFP-----TOLKELFRILISDTOTPTTKRIAL----- 982
DB 818 -LPQFPPDRPGAPTPGSVDKQAIINDLMQLTADSSVPVPAQAQKALCNSGSSFNPNRPG 876
OY 983 ---RL----- 985

```

```

DB 877 QLGRLPYQNLPLDITLQSPGAGPFPIRNSPSYVIPQFGMNGNQMILSGNLGNS 936
OY 986 TDLANTYKSSDPPS-----DOSACERTVLIKLAQLAADQKSMELRSQARCLVALYNINT 1041
DB 937 TGMIGSSTRSMBSMGWAPQSTCESTVLLPLVPRDQKEAR-----FCN 984
OY 1042 PWTLLDLPLKVIQDSARSCITHMRROQSCNSGAN-----SPSSPLSSSS 1090
DB 985 P-----TASIM-----GANSOLQORQMOSQVNMVIGSELEMMNGPQYNOQAAPPNOTA 1035
OY 1091 PKPLSPSPVGPF--ASLOSHHOLISG-----TSPRASQSVQ----- 1128
DB 1036 FWP--ESTLPIDASFASQQRQPPGSDPDLCPHPAESPDSGALLDQLYALRNF 1092
OY 1129 -----ELFSGE-LDI-QHNIQKTS---EEIRHCFGQY--QTALAPNGFN-- 1167
DB 1093 GLEIIDLALGIPELVSGQADABQFSSQESSIMLEQRPVFPQYASQAQAQGVNPM 1152
OY 1168 GHLQYHDQGDQSCASLSSNKTOSANTQSTNPESATMELDN--LEBERTTONAKSPD 1226
DB 1153 QDPNHTWGRPNYTTLRMQPRPGLRPTGIYQONPQLRLQLOHRLQAOQRPLMNOIS 1212
OY 1227 DAKYITVSIN--e-----MAENGELIASNMESEV---VRVALTLTKQPVLL 1270
DB 1213 SVSNVNTLRGVPYTOAPINQMLAOROREITLNQHLRQWOOQOVRTLMWRGGQLVNT 1272
OY 1271 QTSLTNLGICIKGNCCELP 1289
DB 1273 PSMVAPAGLPAAMSNPRIP 1291

```

RESULT 10
US-10-984-645-2

```

; Sequence 2, Application US/10984645
; Publication No. US20050244386641
; GENERAL INFORMATION:
; APPLICANT: Habener, Joel
; APPLICANT: Zulewski, Hendrik
; APPLICANT: Abraham, Elizabeth
; APPLICANT: Vallejo, Mario
; TITLE OF INVENTION: METHOD OF TRANSPLANTING IN A MAMMAL AND TREATING DIABETES MELLI
; TITLE OF INVENTION: BY ADMINISTERING A PSEUDO-ISLET LIKE AGGREGATE DIFFERENTIATED
; FILE REFERENCE: 3284/1223
; CURRENT APPLICATION NUMBER: US/10/984, 645
; CURRENT FILING DATE: 2004-11-09
; PRIOR APPLICATION NUMBER: US 09/731,255
; PRIOR FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215,109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/239,880
; PRIOR FILING DATE: 2000-10-06
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-984-645-2

```

Query Match 1.6%; Score 124; DB 1; Length 1618;
Best Local Similarity 19.5%; Pred. No. 1.1;
Matches 163; Conservative 93; Mismatches 331; Indels 250; Gaps 34;

```

OY 526 TLDIAQAL--ERERGGGGGTGTGTG-----TAPETRYVSR----- 564
DB 295 SLEAVATYKTLLEANSRLQTPGGGSKTSLSPQPKLELPQRPRTPEGRGLSLVLSPS 354
OY 565 --GRTPGLQKTPSMMSISAVDTAAQRAKVAQYTLVYR-----GRKPLGPNNS 613
DB 355 LPSPLPATLTPV-----AFUKQOEFLQARTPTPLASTPTPTQAPSP 399

```

```

QY 614 NOASHTGAAGSLPRPLNLSNGGTP-----ATTPGCVTPRPRGRAGVSGOSQPGSST 667
DB 400 VDAEIRADAPLISLQTOGRKQABEPYRAEARVAIPASVLPFGP-----EEBQ----- 447
QY 668 SPSTLRQOYGGIGNYRGATGAIIPKASGI--PRS--TASSRETSPTSGGLMKRSMYS 724
DB 448 -----GGRQASSTQSPEDHSLAPLSPHSSLEAODGSGGSRVSTCRG 494
QY 725 TGAGS--RRTPERNNVPRSPAPARLLAQSREAEHTLVGDQDPDYVSGDYMRSGEMRGR 783
DB 495 EGEQIIMWLVEKETIIEGVVSSLOQEIWEED-----LNK 530
QY 784 KLMGRDESDIDSEASVCSERSFDSYTRGNKNSYLSGSHTRLDWSTORAP--PDDI 840
DB 531 KEI--QDSQVPLEKETIKSGEIEIGSLKLENO-----SHETLERENQECPRSLIEDL 582
QY 841 ETI-----IOPCASTHWEKDGILSTOVLADQKELTO-----QOLKCVL 881
DB 583 ETLKSLERENKRAIKGCGSETSRKR-----GCRQKLPFGKEDTQTLQSLQKQENQELMKSL 638
QY 882 DMFRKMFMDHTKVSLSF-----LDVTETLIVANETSRN-----GSSSCLTRLENKLTGT 932
DB 639 EGNLETFLEPGTENQELVSSLOENLESITALEKEN--QEPRLSPREVQDEAL--RPLTKENQ 696
QY 933 DLNSWMSKIMKTLOVHEHYEFTQQLKELEFRISDSTQPTTKRIALIRPLETDANTY 992
DB 697 EPLRLSD-----ENKEAFRLSEKENQEPKLEEDQSIVRPLETEN 739
QY 993 CKSSFPEDQOACERTYKLAQ-----LAADQKSELNSQASCVLALYNL----- 1039
DB 740 HKSLRSLSEODETLRTLEKETQORRRSLGQEDQTLRPEPEKVDLEPLSLQELIARPLE 799
QY 1040 NTPQMTLLADLPKYVODSARSCISHMRQOSQNSGANSPPSSPLSSSPKPLQSP-- 1097
DB 800 NENQFL-----KALKESVAVAVSLTEILLESLSKAGQENLETKAPETQAPLMTBEE 853
QY 1098 ---SVGPASLQSHHQLSISTSPRSQSVQD-----LPSSELDIQHNIQKTS 1146
DB 854 INKSGNESSRKGNSRTTCVCGSEPRDIQTPRGESGIIETISGMEPFEIRSGVDKES 913
QY 1147 ---EIRHCFGGQVOTALAPNGFNHLOYHQQGQODSCASLSMSKTOSSANTQSNTP 1202
DB 914 QRLNEEENLKGFEYESLR-----SLEEQGE-----LP 943
QY 1203 ESATWRL--DNLERRTQNAKSP-----TDDAKV-----ITVSINMAENGEL 1243
DB 944 QSAVDQVREDIYVEKQELQAESPPGMAVGENQDEALNLRQODGFTGKEVEVEGEL 1000

```

RESULT 11

```

US-11-046-346-1
; Sequence 1, Application US/11046346
; Publication No. US2005025502A1
; GENERAL INFORMATION:
; APPLICANT: Dana Farber Cancer Institute
; TITLE OF INVENTION: Method for Determination and Quantification of Radiation or Genot
; FILE REFERENCE: 7032/2082
; CURRENT APPLICATION NUMBER: US/11/046,346
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: US 60/540380
; PRIOR FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1451
; TYPE: PR
; ORGANISM: Homo sapiens
US-11-046-346-1

```

Query Match 1.6%, Score 123, DB 7, Length 1451;

Best Local Similarity 17.0%; Pred. No. 1.1, Matches 249; Conservative 212; Mismatches 498; Indels 504; Gaps 63;

```

QY 51 IDGIMPLTSGHFKIAQSLSELSBELIKRLQSDNATATYLP--HVDRLODSDTYREK 109
DB 142 IDLIQPAIITLTF--EKLEPEYFE--NRNDEINIRLLIVSQKMLDRVVDGKDLTKI 196
QY 110 AOL--LRDMERHVL--PPQALIDKLATSCFKHNAKVREELQITVNALHEVTOOLS 165
DB 197 MQLISIPENQHDITISKEPILGDS-----QHDVQKESLDLI-----ENTS 240
QY 166 VRVYIPVCAALLG--DPT--VNVREAIQTLVEIYKAVGDRILRPLRMDVPA----- 215
DB 241 LTVETLDVLSLRIDPFLKVRQLVMDKLSI-----RLEDLPVLIKFIH 287
QY 216 ---SKLAMEQKFDVQKQEGLLPSALKNNTNGVGLDEADNIGLRETRTKIKRPLH 270
DB 288 SVTAMDTLEIVSEIREKLDIQHCVLPRLOASQVTKSKGRASSGQSGGSCIIILF 347
QY 271 SAVSSLRPKKNVNDVYGDAGAVTMESFESFEVVPQINTFHAQMDIDYKQVLIISDK 330
DB 348 DVISALIRYEKTISEA----- 363
QY 331 NADWEKEVD-----ALKKIRALLISYHTQPOFAVOLKELISLFPD--ILKEELRSQVIR 384
DB 364 ---WIKALENTASVSEHKVPDLVWL-----FIVSTNTQTKYIDRLNKRIRSGCIG 413
QY 385 EACITIAVMSKTLNKLDAFCWSILBHLINLQNSAKVIASASTALKYIIRY--THAPK 442
DB 414 EQLQSTFSVAYL--VTKDMCSSTLSLAOSLHSLDQSIISFGSLLYKARFPDTYQO 471
QY 443 ---LTKIYDTLQNSKXDIRSTLCELMVLLFBEWQKALERNATVLRDTLKKSIGDAD 498
DB 472 EVGALVTHICSGEAEEDDLAVLELVV-----NPSAMMANAVFQGL----- 518
QY 499 CDARRHSRYAYMARFRHPELADQIYGLDIAQALREREREGGGGTGTGTAPETR 558
DB 519 -----DYLQNI-----SPOOI 529
QY 559 RTVSRIGTPEGLQKTPPSMSISAVDTAAQAKVRAQVLYLSQRKPLGN--NSNQA 616
DB 530 RKLFFV-----LSTLAFSKONEASHIQQDMHLVIRKOLSTVEKYKI 573
QY 617 SMTGAAGSGLPRRLNSNGGTPATTPGVTLPBPRGAGVSOQ-----PG 663
DB 574 GITAVTMAGI-----MAADRESPSLT-----QERANISDECTQVTSILOLVHGS 621
QY 664 SRSTSPSTKLADQYCGIGNYRGATGAIIPKASGI--PRSTASSRETSPTSGGLMKRSMY 723
DB 622 EQSPQASALYYDEFANLIGHK-----LDPKLBMVGHITCNDFODAFV----- 665
QY 724 STGASRRTPERNNVPRSPAPARLLAQSREAEHTLVGDQDPDYVSGDYMRSG--M 779
DB 666 ---VDSQVPEGDPF-----PVKALYGLEEYDQDGININLPLFLFSQDFADGGPVTQ 718
QY 780 RMGRKLMGRDSDIDSEASVCSERSFDSYTRGNKNSYLSGSHTRLDWSTORAPFDD 839
DB 719 ESGGLTV-----SPIC-----LAPY-- 733
QY 840 IETIIOFCASTHWS--ERKDLISLTOYLD--GKELTOOQLKVLDMFRKMFMDHT 893
DB 734 -FRLRLRVERQHNQNLNLEIDQLDCPIFLDLEBGEKLESMAKASAFMCSLIF----- 787
QY 894 KYVSLFLDTVELLIVHANEYSRNGSSCLTRLPKFKGTDLLNSHNSKIKWTLQVHEVF 953
DB 788 ---LTLNMFREIYVAFQCEITSPEKNGKVLRLKIVELQIL-----LEKTLAVTPDV 837
QY 954 PT-----QOLKELEFRISDSTQPTTKRIAL-----LRFYTDLANTYCKSSDFPS 1000
DB 838 PPLGNFDEVTLDI-----TPHTVTAISAKIRKKGIERKQVTDGSKT--SSSDTIS 886
QY 1001 DOSQA--CERTVUKLAQLAAD-----OKSMELRSQASCLVALYNL 1039

```

```

Db      887 EEKSECDPTSHKGLNKEFTGKEKTSLLHNSHAFRELDIEVSIHCGVTKIL 946
Qy      1040 NT-----PQMTLLADLPKYODSARSCISHMRQSQSCGANSPPSSSP 1085
Db      947 DTEHTEATEVQGLPPELLFLLEDL-----SOKLEMTLPPPIARR 987
Qy      1086 LSSSPKPLQSPSVGPPFASLOSHHHQLSISTSPRSQSSVEQELLFSELDIDHNIQKT 1145
Db      988 V-----PFLKNKG-----SRNIGFSLQORS 1008
Qy      1146 SEERHCGGQGYQTALPANGFNGHLQ-YHDQGGQDSCASLSNSKTSQANTTOS-NTPE 1203
Db      1009 AOEIVHC-----VEQLLTP-MCNLEHNIHNTIQ--CLAENHGVNDGPGYKVOEYHMS 1059
Qy      1204 SATRLDNLERERTQNAKSPPTDAKVTIVSINNAENGLILASNLMESEVVRVALTLTK 1263
Db      1060 SCYRLRLQIFHGLFAMSGFSGPENQNLISALH-----VLSRLKGE-----H 1103
Qy      1264 DQPV-ELLQTS---LTNLGICIKGNCGLPNKHFRSIRMLNLTIE---AETDVIYA 1314
Db      1104 SQPLEELLQSOSVHYLQNFHOSIPSFQCAL-----YLIRLMVILEKSTASQNKKEKIAS 1157
Qy      1315 -GLHVLKSKMS-----NKMHNMMHFLLELILKTIQCYHSHKALRDISMIRIAPSL 1368
Db      1158 LARQFLCRWPSGDKKSKSNISNDLHALCTYL-----EHTESILKALIEIAQVAPEL 1211
Qy      1369 ---PLDLSINIVNPVIATGEFEPT 1388
Db      1212 INSPKXAS-----SSTFPT 1225

```

```

RESULT 12
US-11-074-176-134
; Sequence 134, Application US/11074176
; Publication No. US20050250135A1
; GENERAL INFORMATION:
; APPLICANT: Kluenhammer, Todd R.
; APPLICANT: Russell, William M.
; APPLICANT: Altermann, Eric
; APPLICANT: McAuliffe, Olivia
; APPLICANT: Perill, Andrea Azcarate
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding
; TITLE OF INVENTION: Stress-Related Proteins and Uses Therefore
; FILE REFERENCE: 5051-694
; CURRENT APPLICATION NUMBER: US/11/074,176
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: 60/551,161
; PRIOR FILING DATE: 2004-03-08
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Lactobacillus acidophilus
US-11-074-176-134

```

```

Query March 1.6%; Score 122; DB 7; Length 1189;
Best Local Similarity 18.4%; Pred. No. 0.93;
Matches 209; Conservative 194; Mismatches 447; Indels 286; Gaps 52;

```

```

Qy      134 TSCRHAKKAKREEFLQ-TYNALHEVGTQQLSV---RVYIPYCALIGDPTVNVREAI 189
Db      81 TLIFDNKRELAFTDQVSVTRILRSQDSEFLINNQVWRMDVTTFLDGISPNLSAI 140
Qy      190 QTLVEIYHVGDRL---RPDLRRMDVPAASKLAMEKQFDQVKGGLLPALKNTNGNG 246
Db      141 ISQGRV---DQILNSRPBQRRIIFERAGVLFHKQKKEAQNQ-----LKTQDNL 188
Qy      247 VGLDEADNIGLREPTRMKRPLSHAVSSSLRPKNVNDVTGDGAVTMESSFEVVP 306
Db      189 IRINDL---VKLESRL--EPLNE--QSLAKEVKFPKSGLDKSLALAF-----IE 235
Qy      307 QLN-----IFHAKMDDIYKQVLYITISDNKADWEKRVDAKKIR-ALLILS 351

```

```

Db      236 NINQOREDIQKADKNKILAK--LDDEVKDSQAANVQKREYVKLRBERDHTQKKLKS 294
Qy      352 YHTQPFVAVOLKELSLSPFDILKEELRSQVIREACITIAVMSKTLRNKLDAPFCMSILEH 411
Db      295 KDLSELNASTQMAEQSQFDDATKEBYKNQV-----KQLNQ----- 330
Qy      412 LINLIQNSAKVIAASTIALKYIITKHAPRLKI---YTDILNQ-----SKQIDRS 461
Db      331 -NLVQLKADL-----DELKKEKKQLQDEODVYKIERGQLTGELENDEPEELNKLDIDIRN 383
Qy      462 TLCEMWLLEPEEMQTKALERNATVLRDLTKKSIGDADCDARRHRYVMAFRHFPFLAD 521
Db      384 NYQQL-----ODQFTNNQIV-----NLMSDKRSQ-----AD 412
Qy      522 QIYGTDLIA-----AQRLEREREGGGGTGTGTCTAPETRTVRSIRGPTGLQKPTP 576
Db      413 TTYQGVSVSKQLDADQKQLQRLIEG-----KULTDRQKEQNAIVRINKQNNQNLTELT 467
Qy      577 SMSISAVDTAAQAKAVRAQY-TLYSRQKPLG-----PNSNQ-ASMTGAASGS 626
Db      468 NLRQV--VNAERNELEKVEARHEALVNIQKHREGYVYGVANNVNLHNDPFAVGIVAV--GE 523
Qy      627 LPRRLNNSGCTPATTPGSVTPRPRAGVSSQSPGSRSTPS--TKLDQYQIGINNY 684
Db      524 LIT-----PPAELEAMTTALGG--GVQDLITERSISARAINKLKQNHGGRAF-- 571
Qy      685 RGATGAIIP---KXASGIPRSTASSRSTSPRSQGGKMRSMYSGASRRTPERNNPVR 741
Db      572 -----LPDLGLAQYGIPOSTVTT-----LKSVDGF--RGIASDLVBSKTDQDITAAIN 618
Qy      742 SAPRLLAQRERAEHTLG-----VGDDQDPYVSGDYMRSGGMGRK--LMGRDE 790
Db      619 LLSGVVIVDTIDTAMSAQAQVNRRIYVTLTG--DIVISPGSGMTGGQGNQNSNPLQATE 676
Qy      791 SDDIDSASSVCERSFDS-----YTRGNKSNVSLSGSTRLDMSQRPAPFDIEITIQ 845
Db      677 INOLEKQIKTKLKNLEDDQKLENLYDOSKNVNAELDODALRETSQ-----AIN 727
Qy      846 FCASTHWSERKDGILSTQYIADGKELTQOOLK---CVLDMFRMFMDHTTKVYSLFLDT 902
Db      728 EALISFGQKE-----VKRLDANTLYKSRINDRNRIRELKKQIIEANDK----- 774
Qy      903 VTELILVHNETSRNGSSCLTRL--FNKLGTDLANSMSKIKTKLOYV----- 949
Db      775 --QMLTKQGEEOKAKKNDLQDXIKNFNNLSORIOBEL-SKLPKIAVYTNKLENLSQE 831
Qy      950 -----HEFFPQQLKEL---FRISDSOTPTTKTRIAILRFLDLANTYCKSSDFP--- 999
Db      832 NEKNHQIDNSEKQIEDTAVLTLLAONDENSNQOT-----ANLEKQSTIDQKN 880
Qy      1000 -----SDQSACERTVTLKLAQADQKSMELRSQAR-----SCVALYNLNTPO 1043
Db      881 NELQARLNDLSQLQGDQADQINQLDQVAS--RNYDLRDKDALIEQEDSVYIAKNSSINQ 938
Qy      1044 MTLILADLPKYODSARSCISHMRQSQSCGANSPPSSSPKPLQSPSVGPFA 1103
Db      939 RLEFLRDYSLTFREA--IAQAGEENNEETRNELAKSVKLHRMSIEDIGFVNLDSIOEYE 996
Qy      1104 SLOSHHHQLSISTSPRSQSSVEQELLFSELDIDHNIQKTSEIRHCGGQYQT 1159
Db      997 DVKQRYDFL-----NQQNDLLFARDLEKSMTELDDEVKTRFPGHTPT 1040

```

```

RESULT 13
US-10-485-517-141
; Sequence 141, Application US/10485517
; Publication No. US20050256299A1
; GENERAL INFORMATION:
; APPLICANT: University of Sheffield
; APPLICANT: Biosynexus Incorporated
; APPLICANT: Foster, Simon
; APPLICANT: Mond, James

```

```

; TITLE OF INVENTION: Antigenic Polypeptides
; FILE REFERENCE: P100629WO
; CURRENT APPLICATION NUMBER: US/10/485,517
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: GB 0118825.9
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: GB 0200349.9
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 424
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 141
; LENGTH: 1290
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-10-485-517-141

Query Match      1.6%; Score 121.5; DB 1; Length 1290;
Best Local Similarity 17.6%; Pred. No. 1.1;
Matches 242; Conservative 186; Mismatches 510; Indels 435; Gaps 56;

OY 51 IDGLMPULGSHFKIAQKSLFAFSELIRKLSDFNAATATVLPVHIDRLGDSRDTVREKA 110
DB 152 IDGL-----SHLTNAQK--EHLKQVVOGSTT-----VAEAGN-----182
OY 111 QLLRLDMEHRYLPFOALIDKLATSCEPKHAKAKVEEFLQTVLVAHEYGTOQLSVRYI 170
DB 183 -----EOKANNVDAAMDRLRQSIADNATTKONQNTDASQNKKDAYNNNAVTTAQGI 234
OY 171 PIVCALLGDPYNNVREAIQTLVEIKYKVRGLRDLARMDDVPASKLAML--EQKFDQVK 229
DB 235 DQTSPTLDPVTV--INQAAQO-----VSTTKALNGENLELAK 271
OY 230 QEGLLPSALKNTNG-----NGV--GLDEADNIGLERPRTMIRPLHSAYSSLR 278
DB 272 QOASQSLGSLDYNNAQOQOTVTDQNGAHTYDEANOIQOMQONLNTANGNLKQALADDA 331
OY 279 PKPVNDVTDG-----AGAVTMESEFSEFEV-----POL 308
DB 332 TKATNPTFDADQAKQOAVNTAVTNAENIISKANGNATQAEVEQAIKQVNAKQALNGNA 391
OY 309 NIFNAKD-----MDDIYQOVLVI 326
DB 392 NVQAKDEBATALINSSNDLNOAKQKALQOYONATVAGVNNVQKTOAELNNAMTOLKOG 451
OY 327 ISDK-----NADMEKR--VDALKIRALLILSVHPOPOVAVOLKEL--SLSEV 371
DB 452 IADKEQYADGNFVNADDPKQVNAQVAKAEAL-----SATPVVVTPEITLAKNV 506
OY 372 DILKEELR-----SQVIREACITIAVMSKTLRNKLD--AFCSILEHLINLIONSAY 422
DB 507 TOAKNDLNGNTMLATAKQNVQAHIQPLNLNOQDEVSQKTOATLVVNNVAIQAAAT 566
OY 423 IASAST-----IALKYIITK-----YTHAPKLKIYTT-----DTLN 452
DB 567 IINDAMTOLKQGIANKRAQIKGSENVHADDTKQATVADNAVTKAEELIKOTTNPTMPTNIQ 626
OY 453 QSKSDINSTLCELWVLEFEWQTKALEENATVLDLTKSGIDGADCDARHSRYAYNAF 512
DB 627 QALTK--VNDT-----NQLNLGN-----QKADAKQDK--653
OY 513 RHHPFLADQIYGTLD--IAAORALEREREGGGGTGTGTGTAPETRRVYSRIGRTP--568
DB 654 -----TTLGTIDHLNDQKQALTTQVE-----QAPDI--ATVNNVQVQNO 691
OY 569 -----GTLQKPTPSMRGSIQAVDTAAQAPAKRQATYLSKQKRLGNNNSQNSM 618
DB 692 NINAMTNINNALQDTEETLINSINFID--ADQAKKDAYTNAVSAHEGILSKANGSNAQ 748
OY 619 T--GAAGGSLPRPLNSGSGTPTATPG-----SVTPRPRGRAGVSGSQPG 663
DB 749 TVEQAMQORVNEKQALNGNDVQKADAKQVITIANDLNOAMTOLKGIADKQOTKAN 808
OY 664 SRSTPSYTLRDOYGGIGNYRGATGATPKKASGIPRSTASSRETSPT-----RS 713

```

```

DB 809 GNFVNADTDKQ-----NAYNNAVVAHEQIISGTPIANVDPOQVQALQVNAQKGLN 861
OY 714 GGGIMKMSWYSTAGSRPRPERNNPVPSAPARLLAQSREAEHTLGVGDGQPPYVSGDY 773
DB 862 GNNHLQVAKDNATATLIDQLEPNLQPKTA-----LKQVSHAEVTVGNAIKQ-----NADA 913
OY 774 MRSGMRMGKMLGRDE--SDIDSEASVSCERSFDSSTYRGKNSVSLSGSHRLDMST 832
DB 914 LNNAMGTLKQOIQANSQVPSVPTQADQKQOQVYNNNA--ANOQOIANGPV-----964
OY 833 QRAPFDIETITIOFCASTHWSERKGLISLTOYLADQKELTQOQKLCVLMFRKMFMDTH 892
DB 965 ---PULTPDVTQ--ANTTNQAKDAL--NDEKLAQK-----QELANLDTLR-----DLN 1010
OY 893 TKVYSLFDVTVELILVHANETSRNNGSSCLTRLEN--KLGTDLNSHAKIKWTLQVYHE 951
DB 1011 QPORDALRNQINQALATVETQKQANQNVNTAMSNKQGIANKDTV-----KASENVHD 1065
OY 952 -----YPTQQLKELFRISDSYOTPTTTRIALIFLTDLANITYCKSSDPSPSQS 1003
DB 1066 ADADKQATYNNVAVSQAGIINQTTNPTLNPDELTR--ALTQVDAKN-----1110
OY 1004 QACERTVLKLAQLAADQKSMELRSQARSCLVALYNLNTPOWTLV-----LADLPKV 1054
DB 1111 -----GLNGEKLATEKONAKADVSGMTHLNDQKQALQKQIDQSPELATVQV 1159
OY 1055 YQDSARSCISHMRQOSQNSGANSPPSSPLSSSFKPLQSPSVGPAPASIQSHHQLST 1114
DB 1160 KQTN--TSLQAMDQSLQALINDKQKQTLADGNYLVADP-----DKQNAVYKQAVA 1205
OY 1115 SSTSFRQSSVEBELFSEELDQHNIOKTSERIRHCFQGYQYOTALAPNGFNHLOYHD 1174
DB 1206 KAEALLNKQSGTN-----EVOAQVESITNENVN-----NAKQALNGN-----1241
OY 1175 QGQODSCASLSNSKTSQNSANTQSNTPESATMELD-----NLERERTYONAKS 1223
DB 1242 ---DNLNAAKQAK--QQLANLTLHNDQKQSPSQITQAPLVYDVTITINKA 1289

RESULT 14
US-10-499-715-2
; Sequence 2, Application US/10499715
; Publication No. US20050250717A1
; GENERAL INFORMATION:
; APPLICANT: BENAYAHU, Dafna
; TITLE OF INVENTION: CHROMATIN REMODELING PROTEIN AS A MARKER EXPRESSED BY STROMAL
; FILE REFERENCE: BENAYAHU-1.1 PCT
; CURRENT FILING DATE: 2004-06-21
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 2897
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-499-715-2

Query Match      1.6%; Score 119; DB 1; Length 2897;
Best Local Similarity 17.9%; Pred. No. 5.3;
Matches 110; Conservative 77; Mismatches 222; Indels 206; Gaps 20;

OY 724 ERNNVPRSPADARLLAQSREAEHTLGVDDQCPDYVSGDYNRSGMRGKRLMGDESD 793
DB 1693 EKXNTIRADPALCFLEIRVGKDEKAVAAEQANDYMGD--VEDPEYKAPAIKFDIDDD 1751
OY 794 IDSEASVCSERSFDSSTYRGKNSVSLSGSHRLDMSTQAPFDIETITIOFCASTH--851
DB 1752 VSSPDLVIADG--DQQLMEGDK-----VYPTQSLATYTRRLRLITAYQRTKAN 1798
OY 852 -----WSEBKD-----857

```

```

Db      1799 RQIQOIPPTSVPTSVNQPPIYEATLNPXAKIEROQRTREAEFYRVVSTFGVFD 1858
Qy      858 ---GLISLTOYLADOK-----ELTQOOLKCVLDMFRKM-----FMDHTHTVYSLFL 900
Db      1859 PDRQOPMTKRAARLHKKTDDSLKLYAFMCMCRVCLPSKEELVD-----NIFI 1913
Qy      901 DVTLELLVHANETSRNGSSSCLTRLFNKLTGTDLLNSMHSXIMKTLOVH-----E 951
Db      1914 QPIRE-----ERASRTLYRIELRKVREQALR--HPQLFRLKLCHEPNPLPWME 1962
Qy      952 YFPFQOLQ-----KELFRIISD-----STQPTTKRTAIALFLDLANTY--- 992
Db      1963 CGPHDRDLIGAAGKVSRTYHILRDELSFMAAQRYSQSKVAHSRTSPILQOQVA 2022
Qy      993 -----CKSADFPSSDOQACE-----RTVVK----- 1012
Db      2023 LSASPLTSLPRLAKGILBEMVKSENLEKEFQSSSEBSMSVETRTLIKSEPVAPKN 2082
Qy      1013 --LAQLADOKSMELRSQARSCLVALYNLTPOMTLLADLPKYODSARSCISHNRQ 1070
Db      2083 GVLPQATGDKSGKGCETDRMVAARTEPLTPNPA---SKKPRVHKRGSESSSDSDSE 2139
Qy      1071 SQSNGSANGSPSSPSSSPKPLQSPSVGPFALOSHHQLSISSTSPRSQSSVQEL 1130
Db      2140 RSSCSRSSSSSSSSSHSRSGSSSSSSSSCSASSS---SSSSTSSSSSSSSSSSEE 2195
Qy      1131 LFSSELDIHNHIOKTSSEIRHCFGGQYQTALAPNGFNHLOYHGGODSCASLSNSKT 1190
Db      2196 SDSDEEAQKRETT-----HMKAYD---EBSVASLSTTQDE 2229
Qy      1191 QSSANTQOSNTPESA 1205
Db      2230 TQDSFQMNNGTPESA 2244

RESULT 15
US-10-821-234-1070
; Sequence 1070, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821,234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ ID NOS: 1704
; SOFTWARE: PL_SEQ_genes version 1.0
; SEQ ID NO 1070
; LENGTH: 1377
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1070

```

```

Query Match      1.5%; Score 116.5; DB 1; Length 1377;
Best Local Similarity 21.3%; Pred. No. 2.7;
Matches 115; Conservative 46; Mismatches 198; Indels 181; Gaps 24;

```

```

Qy      456 SKDRLTLCEMLVLLFEEMQTKALERNATVLRDTLKKSIGDAD-----DARRHRYAY 509
Db      320 AQNHETITLBMKALKELFEFRPSKS-----QSSGSAHPISVPGARRH----- 365
Qy      510 WAFRRHPEEL-----ADQIYGLDIAQALEREREGGGGTGTGTAPET 557
Db      366 ---HHLVNLPPSGTGLVRSRRTSLAAT--PPAAKCSSCRVRTASEGDG--GAAAGAAAA 419
Qy      558 RRTYSRIGR--TPGTLQKPTSMRSISAVDTAAAGRAKVAQYTLYSRQRKPLGPNNSNQ 615

```

```

Db      420 QRLVSVAGSEPLSGPVAPARLSRHTLSG---GCCGRGSKVAL-----LPAGALQHS 468
Qy      616 ASMT-----GAAAGSLPRPRLNSNG---GTPATTPGSVTPPR-----GRAGVS 658
Db      469 RSMSPVVAHSPPAATSPGS-----LSSSGHSGSYRPPPGPHPLPHLHGGPGGPPSS 523
Qy      659 QSQPGSRSTS--PSTKLRDQYGGIYRQATGAIIPKASGIPSTASRRTSPTRSGGG-- 716
Db      524 GSASAGSPSPDPGFMGLDEYGS---SPGDLRAFCSHRSNTPEISA---ETPPARDGGGG 576
Qy      717 -----LMKRSWYSTGAGSR-----TPRRNPVR----- 740
Db      577 GERYGYMTDRPLSHCGRSYRVRVSGDAODLRGLAKRTYSLTPPARQRPVPPSSASLD 636
Qy      741 -----PSAPARL-----LAQSRBAEHTLGVGDGQPD 767
Db      637 EYTLMRATSGSAGRLCPGCPASSPKVAHYRPEDYDIEIGSHRSSSSSLGADDGYPM 696
Qy      768 YVSGDYMRGGGMKMKLMGRDESD-----IDSEASSVCERSF 807
Db      697 TPGAFAFGSGS-----GSCRSDDYMEMSPASVAPKQILOPRAAAAAAIVPAGPA 748
Qy      808 DSGYTERGKNSNYLSGSHTRLDWSTORAPEDDIETIIOPCASTHMGSRKDG--LISLTOYL 866
Db      749 GPAPTAAAGRTFPASGGYKASSPAESSP--EDSGYKRMCGSLSMHADGKLLPNGDYL 807

```

Search completed: November 23, 2005, 15:17:09
Job time : 25 secs

Wed Nov 30 10:28:06 2005

us-10-030-850-1.rge

10/630850 Page 1
See IDS 1 & 2

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: November 29, 2005, 14:18:08 ; Search time 16694 Seconds
(without alignments)
17518.843 Million cell updates/sec

Title:	US-10-030-850-1
Perfect score:	5145
Sequence:	1 gaattcgcgacgagcgctc.....ggatcaaaatcataaaaaa 5145

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

```
Searched:      5883141 seqs, 28421725653 residues
Total number of hits satisfying chosen parameters: 11766282
```

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

No.	Score	Query Match	Length	DB	ID	Description
1	5145	100.0	5145	6	AX073964	AX073964 Sequence
2	5120.8	99.5	5959	2	AB031048	AB031048 Drosophila
3	5055.6	98.3	5916	2	AF195498	AF195498 Drosophila
4	5055.6	98.3	5938	2	AF250842	AF250842 Drosophila
5	5055.6	98.3	5962	2	AY065579	AY065579 Drosophila
6	5051.8	98.2	87197	14	AC014071	AC014071 Drosophila
7	5051.8	98.2	163098	2	AC010055	AC010055 Drosophila
8	5051.8	98.2	182960	2	AC010017	AC010017 Drosophila
9	5051.8	98.2	313157	2	AE003593	AE003593 Drosophila
10	5051.8	98.2	313157	2	AE003593	AE003593 Drosophila
11	1824.4	37.4	96091	14	AC130040	AC130040 Rattus norvegicus
12	1824.4	35.7	79733	14	AC006579	AC006579 Drosophila
13	732.8	14.2	767	6	AR495636	AR495636 Sequence
14	732.8	14.2	767	6	AR510918	AR510918 Sequence
15	661	12.8	1420	6	AR501058	AR501058 Sequence
16	661	12.8	1420	6	AR516340	AR516340 Sequence
17	452.4	8.8	466	6	AR509306	AR509306 Sequence
18	331.4	6.4	339	10	G01355	G01355 fruit fly
19	149.8	2.9	3153	6	AX746802	AX746802 Sequence

19	19.8	2.9	3153	8	AK091140
20	19.8	2.9	7888	8	C0873990
21	19.8	2.9	7888	8	AF347693
22	19.4	2.9	154	6	AR525441
23	141.8	2.8	2524	6	AX746992
24	141.8	2.8	2524	8	AK091451
25	140.2	2.7	7426	5	D0022563
26	125.6	2.4	4855	5	BC066436
27	115.6	2.3	1583	8	BC029035
28	116.6	2.3	1908	8	AX746588
29	116.6	2.3	1908	8	AK090716
30	115.6	2.3	5614	6	AB683140
31	115.6	2.3	5614	8	AB014527
32	115.6	2.3	6445	6	AR336833
33	115.6	2.3	6487	6	AR336832
34	112.2	2.2	4494	9	BC039216
35	107.6	2.1	5570	9	RNC028806
36	107	2.1	1268	9	MMU276962
37	104.4	2.0	5191	9	AB093252
38	102.8	2.0	3098	8	AB014542
39	98.4	1.9	6951	8	AB014522
40	94.2	1.8	1189	8	HS2488057
41	92	1.8	608	6	C0527362
42	90	1.7	627	5	BX333073
43	89.8	1.7	5770	8	HSN809084
44	86.6	1.7	5906	8	AK074338
45	81.8	1.6	3141	2	AK116783
					AK091140 Homo sapi
					C0873990 Sequence
					AR347693 Homo sapi
					AR525441 Sequence
					AX746992 Sequence
					AK091451 Homo sapi
					D0022563 Xenopus l
					BC066436 Datto rer
					BC029035 Homo sapi
					AX746588 Sequence
					AK090716 Homo sapi
					AB683140 Sequence
					AB014527 Homo sapi
					AR336833 Sequence
					AR336832 Sequence
					BC039216 Mus muscu
					AJ788860 Rattus no
					AJ767662 Mus muscu
					AB093252 Mus muscu
					C0715342 Sequence
					AB014522 Homo sapi
					AJ288057 Homo sapi
					C0527362 Sequence
					BX333073 Gallus ga
					BX648933 Homo sapi
					AK074338 Homo sapi
					AK116783 Clona int

ALIGNMENTS

	RESULT	1			
	LOCUS	AX073964	5145 bp	DNA	linear
	DEFINITION	Sequence 1 from Patent WO0104295.			
	ACCESSION	AX073964			
	VERSION	AX073964.1	GI:12710224		
	SOURCE				
	KEYWORDS				
	ORGANISM	<i>Drosophila melanogaster</i> (fruit fly)			
		<i>Drosophila melanogaster</i>			
		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
		Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
		Epitheroidea; Drosophilidae; <i>Drosophila</i> .			
REFERENCE	1 Avides,M.D., Deak,P. and Glover,D.M. Orbit and homologues thereof Patent: WO 0104295-A 1 18-JAN-2001; University of Dundee (GB) Location/Qualifiers 1..5145 /organism=" <i>Drosophila melanogaster</i> " /mol_type="unassigned DNA" /_db_xref="taxon:7227"				
FEATURES	source				
ORIGIN	Query Match 100.0%; Score 5145; DB 6; Length 5145; Best Local Similarity 100.0%; Pred. No. 0; Matches 5145; Conservative 0; Mismatches 0; Indels 0; Gaps 0				
Gy	1 GAATTGGCAGGAGCCGCTCGAAATTTGTTCGCCGACAAGAACGAAGAATGACTTG 60				
Dd	1 GAATTCGCACGAGCCGCTCGAATAATTTGTTGCCGACAAGAACGAAGAATGACTTGG 60				
Gy	61 CAACGATGGCTTAATCGAAGCCCAAGCACCTTGATGGTTTTATCCAGCAAATGCCCAAG 120				
Dd	61 CAAAGATGGCTTAATCGAAGCCCAAGCACCTTGATGGCTTTATCCAGCAAATGCCCAAG 120				
Gy	121 CGGAATATGCTGTGAAGTACAGCTCGCGCAGCATCTGTGACATTTCTTAGCGAGACA 180				
Dd	121 CGGATATGCTGTGAAGTACAGCTCGCGCAGAGATCTGTGTACATTTCTTAGCGAGACA 180				
Gy	181 CAACATCAATGTCTGCACGGAATGGAGATTCCTATTAGCGGTTGATGCCATGGCTGA 240				

```
/organism="Drosophila melanogaster"  
/mol_type="unassigned DNA"  
/db_xref="taxon:7227"
```


D	b		181	CAAACTCAATTGTCTGCGACGACATGGAGTTCTTATTGACGGTTTGATGCCATGGCTGA	240
Q	y		241	CGGCGAGCCACTTTAAGATTGGACAAAAGTCCCTGGAGCGTTCTCGAGCTAATPAAAGC	300
D	b		241	CGGCGAGCCACTTTAAGATTGGACAAAAGTCCCTGGAGCGTTCTCGAGCTAATPAAAGC	300
Q	y		301	GATTGGGCGAGCGATTTTAAATGCAATACGCGGCTACCGGTTCTGCGACATGTGATCGAGC	360
D	b		301	GATTGGGCGAGCGATTTTAAATGCAATACGCGGCTACCGGTTCTGCGACATGTGATCGAGC	360
Q	y		361	TGGGAGACAGCAGGGACACAGTCCGCGAAGAGCGCGCACTTCTGCTGGCGACCTCATGG	420
D	b		361	TGGGAGACAGCAGGGACACAGTCCGCGAAGAGCGCGCACTTCTGCTGGCGACCTCATGG	420
Q	y		421	AGCAGAGAGTGTTCGCGCCCGAGCGCTGATGACAAAGCTGGCCACTAGCTGCTTCAAGC	480
D	b		421	AGCAGAGAGTGTTCGCGCCCGAGCGCTGATGACAAAGCTGGCCACTAGCTGCTTCAAGC	480
Q	y		481	ACAAAGACGCCAAGGTGGCGAGAGATTCTTCAGACGATGTGAAACGCTCTTCATGAGT	540
D	b		481	ACAAAGACGCCAAGGTGGCGAGAGATTCTTCAGACGATGTGAAACGCTCTTCATGAGT	540
Q	y		541	AACGACCCACAGCTTATGTTGCGCGTCTAATPCCACAGTTTGTGCACTTCTCGGAG	600
D	b		541	AACGACCCACAGCTTATGTTGCGCGTCTAATPCCACAGTTTGTGCACTTCTCGGAG	600
Q	y		601	ATCCCAACGTTAATGTGAGGAGGCGGCGCATCCAAACGCTAGTGGAAATCTACAAAGCATG	660
D	b		601	ATCCCAACGTTAATGTGAGGAGGCGGCGCATCCAAACGCTAGTGGAAATCTACAAAGCATG	660
Q	y		661	TAGGGGATCGATTGGCGCCAGACCTCCGTTGCGATGACGATGTTCTTGCGAAATTTGG	720
D	b		661	TAGGGGATCGATTGGCGCCAGACCTCCGTTGCGATGACGATGTTCTTGCGAAATTTGG	720
Q	y		721	CTATGTTGGACCAAAAAGTTTCGACCAAGGTGCAACAGAGAGGTCTACTGTACTTCAAGCC	780
D	b		721	CTATGTTGGACCAAAAAGTTTCGACCAAGGTGCAACAGAGAGGTCTACTGTACTTCAAGCC	780
Q	y		781	TTTAAAAACACCAATGTGCAATGAGTGGGCTTTGACGAGGCGGACAAATTTGGGTTGAGG	840
D	b		781	TTTAAAAACACCAATGTGCAATGAGTGGGCTTTGACGAGGCGGACAAATTTGGGTTGAGG	840
Q	y		841	AGGACCCACAGAGATGTTAAGCGGCGCACTACACTCGGCGGTTTCGTATCACTGCGCC	900
D	b		841	AGGACCCACAGAGATGTTAAGCGGCGCACTACACTCGGCGGTTTCGTATCACTGCGCC	900
Q	y		901	CAAAACCCAAATGTGAACGATGTGACCCGGTGAATGCCGGCCGTTAACCATGGAATCTTTGG	960
D	b		901	CAAAACCCAAATGTGAACGATGTGACCCGGTGAATGCCGGCCGTTAACCATGGAATCTTTGG	960
Q	y		961	AATCTAGCTTGAAGTGGTCCCGCAATTTGGAACATTTTCACAGCTTAAGGACATGAGCAGTA	1020
D	b		961	AATCTAGCTTGAAGTGGTCCCGCAATTTGGAACATTTTCACAGCTTAAGGACATGAGCAGTA	1020
Q	y		1021	TCTTCAAGCAAGTACTAGTGCATCACTAGTGTAAAAAACGACAGCTGGGAGAAAGTGTGG	1080
D	b		1021	TCTTCAAGCAAGTACTAGTGCATCACTAGTGTAAAAAACGACAGCTGGGAGAAAGTGTGG	1080
Q	y		1081	ATGCTCTCAAGAAAGTACAGGCGCATTTGCTCATTTCTAGCTATCACACTAGCGCAGATTGG	1140
D	b		1081	ATGCTCTCAAGAAAGTACAGGCGCATTTGCTCATTTCTAGCTATCACACTAGCGCAGATTGG	1140
Q	y		1141	TTCGCTGTACACTTAAGAAATTTGTCGTTAAACCTTCGTGGAGATCCTTCAAGAGGAACATAC	1200
D	b		1141	TTCGCTGTACACTTAAGAAATTTGTCGTTAAACCTTCGTGGAGATCCTTCAAGAGGAACATAC	1200
Q	y		1201	GATCACAGGTATCGCGAGGCGTGCATCACATCGCTACATGTCTTAAGACGCTGAGAA	1260
D	b		1201	GATCACAGGTATCGCGAGGCGTGCATCACATCGCTACATGTCTTAAGACGCTGAGAA	1260
Q	y		1261	ATTAACCTAGATGCTTCTGCTGAGACATTTTGGAGCACCTGATTAATTAATACAGAAACA	1320
D	b		1261	ATTAACCTAGATGCTTCTGCTGAGACATTTTGGAGCACCTGATTAATTAATACAGAAACA	1320

Db	1261	ATAAAGTAAAGCTCTGCTGCGAGCAATTTGGAGCACTGATTAATTAAATACAGAA	1320
Qy	1321	GCGGAGAGTCATTGCAATCCGGCTTCCAAATAGCTTGAAAGTATATCAATTAAATAC	1380
Db	1331	GCGGAGAGGTCAATGCAATCCGCTTCCAAATAGCTTGAAAGTATATCAATTAAATAC	1380
Qy	1381	ATGCAACCAABAGTCTCAAGATCTACACAGACTGTGATCATCAAAGTCGAAGCA	1440
Db	1381	ATGCAACCAABAGTCTCAAGATCTACACAGACTGTGATCATCAAAGTCGAAGCA	1440
Qy	1441	TAAAGTCCACACTGTGTGAGCTGATGGTGTGCTCTTCGAGAGTGGCAGACGAGCGT	1500
Db	1441	TAAAGTCCACACTGTGTGAGCTGATGGTGTGCTCTTCGAGAGTGGCAGACGAGCGT	1500
Qy	1501	TGAAAGGAATGCCACGCTACTAAAGGACACTTAAATAATCATTTGGCGATCAACT	1560
Db	1501	TGAAAGGAATGCCACGCTACTAAAGGACACTTAAATAATCATTTGGCGATCAACT	1560
Qy	1551	GCGATGCAACGCGCGCATTTCCAGATAGCGCTATTTGGGCTTCAAGCGTCACTTCCBAGC	1620
Db	1551	GCGATGCAACGCGCGCATTTCCAGATAGCGCTATTTGGGCTTCAAGCGTCACTTCCBAGC	1620
Qy	1621	TGGGGGATCAAAATATATGGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAAC	1680
Db	1621	TGGGGGATCAAAATATATGGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAAC	1680
Qy	1681	GAGAGGGCGCGGAGAGAGAACTGTATCTGGGACTGGGACTGCACCTTGAAAGAAC	1740
Db	1681	GAGAGGGCGCGGAGAGAGAACTGTATCTGGGACTGGGACTGCACCTTGAAAGAAC	1740
Qy	1741	GCACTGTATCCCGCATTTGGCCGAACACTTGGAACCTTCGAAAAAGCCACGCTTATGTA	1800
Db	1741	GCACTGTATCCCGCATTTGGCCGAACACTTGGAACCTTCGAAAAAGCCACGCTTATGTA	1800
Qy	1801	GATCATTTCGAGCGGTGGAACATCGCGCTGCTCAACAGAGCAAAAGTTAGAGCGCAATATA	1860
Db	1801	GATCATTTCGAGCGGTGGAACATCGCGCTGCTCAACAGAGCAAAAGTTAGAGCGCAATATA	1860
Qy	1861	CACATATATTCGAGCAAAAGAAACCTTTAGAACCTTAATAATTCCAACGAGCATGATGA	1920
Db	1861	CACATATATTCGAGCAAAAGAAACCTTTAGAACCTTAATAATTCCAACGAGCATGATGA	1920
Qy	1921	CAGAGACGACGACGATGGAATCACTGCGCCAGACTCGCTGGAATTCCAATAGCGGTGGCA	1980
Db	1921	CAGAGACGACGACGATGGAATCACTGCGCCAGACTCGCTGGAATTCCAATAGCGGTGGCA	1980
Qy	1981	CACGACGTACACGCGCGGATTCGGTTACACACGCGCCGCGGAGCGGAGAGATATCTC	2040
Db	1981	CACGACGTACACGCGCGGATTCGGTTACACACGCGCCGCGGAGCGGAGAGATATCTC	2040
Qy	2041	AGTCACAGCAGATCGAGATCCACCTCGCCAAAGCAAAAAGTGAGGATCAGTACGGTG	2100
Db	2041	AGTCACAGCAGATCGAGATCCACCTCGCCAAAGCAAAAAGTGAGGATCAGTACGGTG	2100
Qy	2101	GTATTGGTAATTACTATCGTGGGGCCACTGCGCATATCCCAAAAAGGCTTCGGGAATAC	2160
Db	2101	GTATTGGTAATTACTATCGTGGGGCCACTGCGCATATCCCAAAAAGGCTTCGGGAATAC	2160
Qy	2161	CCCGAAGCAGCGCACTCTTAGGGAAAACAGATCCACTAGATCAGGTGTGGCTTGAATGA	2220
Db	2161	CCCGAAGCAGCGCACTCTTAGGGAAAACAGATCCACTAGATCAGGTGTGGCTTGAATGA	2220
Qy	2221	AACGCAATATGTAATCTACAGAGTGGGGGATCTCGACTAGCGCCGAGAGAAACAACCGAG	2280
Db	2221	AACGCAATATGTAATCTACAGAGTGGGGGATCTCGACTAGCGCCGAGAGAAACAACCGAG	2280
Qy	2281	TAAAGCATCGGCGCGGACGACTGTGGCGCAATCCGCTGAGAGCAAAATATCATTTAG	2340
Db	2281	TAAAGCATCGGCGCGGACGACTGTGGCGCAATCCGCTGAGAGCAAAATATCATTTAG	2340
Qy	2341	GCGTCGAGATGATGACACGCACTATGTTTCGCGGAGCTACATGCGAGCGCGGAG	2400
Db	2341	GCGTCGAGATGATGACACGCACTATGTTTCGCGGAGCTACATGCGAGCGCGGAG	2400

Db 1130 AGAGTCCTCCGCCAGCGCTGATGACAGCTGGCCACTAGCTGCTTCAAGCAAG 1189
QY 486 AACGCCAAGGTGGCGAGGAGTTCCTTACAGCATTTGTGAACGCTCTTCACTAGTACGG 545
Db 1190 AACGCCAAGGTGGCGAGGAGTTCCTTACAGCATTTGTGAACGCTCTTCACTAGTACGG 1249
QY 546 ACCCAGCAGTTAGTGTTCGCTCTATATACACAGTTGTGCACTTCTCGAGATCCC 605
Db 1250 ACCCAGCAGTTAGTGTTCGCTCTATATACACAGTTGTGCACTTCTCGAGATCCC 1309
QY 606 ACAATTATGTGAGGAGCGCGCCATCCAAACCTTAGTGAATCTACAGCATGTAAGG 665
Db 1310 ACAATTATGTGAGGAGCGCGCCATCCAAACCTTAGTGAATCTACAGCATGTAAGG 1369
QY 666 GATCGATTGGGCCCAACCTCCGTCGATGAGATGTTCTCCCTCGAAATTGGGCTATG 725
Db 1370 GATCGATTGGGCCCAACCTCCGTCGATGAGATGTTCTCCCTCGAAATTGGGCTATG 1429
QY 726 TTGGAGCAAAAGTTCCACACAGTCAAAAGAGGGTCTACTGCTACCTTCAGCCTTAA 785
Db 1430 TTGGAGCAAAAGTTCCACACAGTCAAAAGAGGGTCTACTGCTACCTTCAGCCTTAA 1489
QY 786 AACACGAATGCAATGAGTGGGCTTGAAGAGCGCGACAAATTTGGTTGAGGAGCCA 845
Db 1490 AACACGAATGCAATGAGTGGGCTTGAAGAGCGCGACAAATTTGGTTGAGGAGCCA 1549
QY 846 CCCACACAGATGATTTAAAGCGGCCACTACACTCGGCGTTTCGTCATCACTCGGCCCA 905
Db 1550 CCCACACAGATGATTTAAAGCGGCCACTACACTCGGCGTTTCGTCATCACTCGGCCCA 1609
QY 906 CCCAATGTGACGATGTGACCGGTGATGCGGCGCGCTAAACATGGAATCTTTCGATCT 965
Db 1610 CCCAATGTGACGATGTGACCGGTGATGCGGCGCGCTAAACATGGAATCTTTCGATCT 1669
QY 966 AGCTTTGAGTGTCCCGCAATTGAAATCTTCCACGCTAAGACATGACATATCTAC 1025
Db 1670 AGCTTTGAGTGTCCCGCAATTGAAATCTTCCACGCTAAGACATGACATATCTAC 1729
QY 1026 AAGCAATCTACTAGTATCATCACTGATTAATAAGCAGACTGGAGAAAGCTGTGATGCT 1085
Db 1730 AAGCAATCTACTAGTATCATCACTGATTAATAAGCAGACTGGAGAAAGCTGTGATGCT 1789
QY 1086 CTCGAAGAATCAGGACATTTGCTCATTTCTAGCTATCACTAGCGCGCAGTTTGTGCT 1145
Db 1790 CTCGAAGAATCAGGACATTTGCTCATTTCTAGCTATCACTAGCGCGCAGTTTGTGCT 1849
QY 1146 GTACAGCTAAAGAAATTTGCTTAAGCTTGTGGAACATCTCAAGAGAACTACATCA 1205
Db 1850 GTACAGCTAAAGAAATTTGCTTAAGCTTGTGGAACATCTCAAGAGAACTACATCA 1909
QY 1206 CAGGTATCCGAGGCGGTGATCACTATGCTCACTGATTAAGACGCTGAAATTA 1265
Db 1910 CAGGTATCCGAGGCGGTGATCACTATGCTCACTGATTAAGACGCTGAAATTA 1969
QY 1266 CTAGATGCTCTTGTCTGAGACATTTTGAACACCTGATTAATTAATACAGACGCG 1325
Db 1970 CTAGATGCTCTTGTCTGAGACATTTTGAACACCTGATTAATTAATACAGACGCG 2029
QY 1326 AAGGTATGTCATCCGCTTCCACAATAGCTCTGAAGTATCATTAAGTATACATGCA 1385
Db 2030 AAGGTATGTCATCCGCTTCCACAATAGCTCTGAAGTATCATTAAGTATACATGCA 2089
QY 1386 CCAAAGCTGTCAAGATCTACAGACACTGTGAATCAAGTGAAGGACATPAAG 1445
Db 2090 CCAAAGCTGTCTAAGATCTACAGACACTGTGAATCAAGTGAAGGACATPAAG 2149
QY 1446 TCCACACTGTGTAGCTGATGTGCTCTTTCGAGGAGTGGGACGAAGGCGTTGGA 1505
Db 2150 TCCACACTGTGTAGCTGATGTGCTCTTTCGAGGAGTGGGACGAAGGCGTTGGA 2209
QY 1506 AGGAATGCCACCGTACTTAAGGACACTTAAAAAATTCATTGGCGATGCAACTGGAT 1565
Db 2210 AGGAATGCCACCGTACTTAAGGACACTTAAAAAATTCATTGGCGATGCAACTGGAT 2269

QY 1566 GCACGCCGCATTCAGATATACGCTATTGGGCTTTACGGCGTCACTTTCAGAGCTGGCG 1625
Db 2270 GCACGCCGCATTCAGATATACGCTATTGGGCTTTACGGCGTCACTTTCAGAGCTGGCG 2329
QY 1626 GATCAATTTATGGAACATTTAACAATAGTGGCCAGCGGCACTTGAAGGGAACGAGAG 1685
Db 2330 GATCAATTTATGGAACATTTAACAATAGTGGCCAGCGGCACTTGAAGGGAACGAGAG 2389
QY 1686 GCGCGGCGAGGAGGGAACCTGGTACTGGGACTGGGACTGCACCTGAAACGACACT 1745
Db 2390 GCGCGGCGAGGAGGGAACCTGGTACTGGGACTGGGACTGCACCTGAAACGACACT 2449
QY 1746 GTATCCCGCATTTGGCGGAACACTTGAACCTTGCAAAAGCCCAACGCTATGATGATCC 1805
Db 2450 GTATCCCGCATTTGGCGGAACACTTGAACCTTGCAAAAGCCCAACGCTATGATGATCC 2509
QY 1806 ATTTCAAGCGGTGGAACATGCGGCTGTCAACGAGCCAAAGTTAGACGCAATTAACCTA 1865
Db 2510 ATTTCAAGCGGTGGAACATGCGGCTGTCAACGAGCCAAAGTTAGAGCGCAATTAACCTA 2569
QY 1866 TATTCAGGCAAAAGGAACCTTTAGSACCTAATTAATTCACACGAGCATGATGACAGA 1925
Db 2570 TATTCAGGCAAAAGGAACCTTTAGSACCTAATTAATTCACACGAGCATGATGACAGA 2629
QY 1926 GCAGACGCGATGATCACTGCGCAGACCTCGCTGAATTCGAATAGCGGTGCAACCA 1985
Db 2630 GCAGACGCGATGATCACTGCGCAGACCTCGCTGAATTCGAATAGCGGTGCAACCA 2689
QY 1986 GCTTACAACGCGGGATTCGTTTACACACGCGCCCGGGACGGGACGAGATTTCTAGTCA 2045
Db 2690 GCTTACAACGCGGGATTCGTTTACACACGCGCCCGGGACGGGACGAGATTTCTAGTCA 2749
QY 2046 CAGCAGGATGAGATTCACCTCGCCAGACAAACCTGAGGGATCAGTACGGTGTATT 2105
Db 2750 CAGCAGGATGAGATTCACCTCGCCAGACAAACCTGAGGGATCAGTACGGTGTATT 2809
QY 2106 GGTAAATTACTATGTTGGGCGCACTGGCGCATACCCAAAGAGCTTCGGAAATACCCCA 2165
Db 2810 GGTAAATTACTATGTTGGGCGCACTGGCGCATACCCAAAGAGCTTCGGAAATACCCCA 2869
QY 2166 ACGACAGCCAGCTCTAGGGAACGAGTCCAACTAGGTCAAGTGTGGCTTGAATGAACCG 2225
Db 2870 ACGACAGCCAGCTCTAGGGAACGAGTCCAACTAGGTCAAGTGTGGCTTGAATGAACCG 2929
QY 2226 AGTATGTACTCTACAGGTGCGGGGTCTGACGTAGCGCGAGAGGAACAACCACTAAGA 2285
Db 2930 AGTATGTACTCTACAGGTGCGGGGTCTGACGTAGCGCGAGAGGAACAACCACTAAGA 2989
QY 2286 CCATCGCGCGCGCAACCTGCTGGGCAATCCCGTGAAGCAGAACATACATTAGGCGTTC 2345
Db 2990 CCATCGCGCGCGCAACCTGCTGGGCAATCCCGTGAAGCAGAACATACATTAGGCGTTC 3049
QY 2346 GGAATGATGGAACAGCCAGACTATGTTTCCGGGACTACATGCGCAGCGCGGAATGCCG 2405
Db 3050 GGAATGATGGAACAGCCAGACTATGTTTCCGGGACTACATGCGCAGCGCGGAATGCCG 3109
QY 2406 ATGGGTGAAGAACTCATGAGGAGTGAATGATGATGATGATGATGATGATGATGATGAT 2465
Db 3110 ATGGGTGAAGAACTCATGAGGAGTGAATGATGATGATGATGATGATGATGATGATGAT 3169
QY 2466 GTGTGTTTCAAGACGATCTTTCAGCTCAGCTACATGAGGATTAATTCGAATCTCA 2525
Db 3170 GTGTGTTTCAAGACGATCTTTCAGCTCAGCTACATGAGGATTAATTCGAATCTCA 3229
QY 2526 CTTAGCGGAGGCAACCTCGCTTGAACCTGGAAGCAGACGCGCGCATTTGACGACATC 2585
Db 3230 CTTAGCGGAGGCAACCTCGCTTGAACCTGGAAGCAGACGCGCGCATTTGACGACATC 3289
QY 2586 GAGAGCATTAATCAGTTCTGCGCATGACGCACTTGGCTGAAGGAAGATGACCTGATC 2645
Db 3290 GAGAGCATTAATCAGTTCTGCGCATGACGCACTTGGCTGAAGGAAGATGACCTGATC 3349

QY 2646 AGCTCACAGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCGTT 2705
 DB 3350 AGCTCACAGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCGTT 3409
 QY 2706 CTGACATGTTCCCGAAGATGTTCAATGATACACACACAGGTGTACTCTGTTTTG 2765
 DB 3410 CTGACATGTTCCCGAAGATGTTCAATGATACACACACAGGTGTACTCTGTTTTG 3469
 QY 2766 GACACGGTCACCGAATGATCTGGTTGATGCGAAGCACTTACGAAATGCGTTTCA 2825
 DB 3470 GACACGGTCACCGAATGATCTGGTTGATGCGAAGCACTTACGAAATGCGTTTCA 3529
 QY 2826 TCTTGTGAGCGCGGTTGTTCAACAAATTTGGGCAACCGATCTACTCAATTCGATGACAGC 2885
 DB 3530 TCTTGTGAGCGCGGTTGTTCAACAAATTTGGGCAACCGATCTACTCAATTCGATGACAGC 3589
 QY 2886 AAGATTTGGAAGAGCTGATAGAGTGGTCCACGAATTTTCCCGACGAGTTGCACTTAA 2945
 DB 3590 AAGATTTGGAAGAGCTGATAGAGTGGTCCACGAATTTTCCCGACGAGTTGCACTTAA 3649
 QY 2946 GAACTATTTAGAAATCATTCGGATTCTPACTCAACGCGCACTACCAAGACGCGATGCC 3005
 DB 3650 GAACTATTTAGAAATCATTCGGATTCTPACTCAACGCGCACTACCAAGACGCGATGCC 3709
 QY 3006 ATTCTGCGCTTCTTACGGATCTGGCCAAATACATCTGTAAAGCAGCGACTTTCCAGC 3065
 DB 3710 ATTCTGCGCTTCTTACGGATCTGGCCAAATACATCTGTAAAGCAGCGACTTTCCAGC 3769
 QY 3066 GACCAAGCCAGGCTCGGAGAGCGATCCTAAAGCTGGCCGACGTGGCGGCGATCAG 3125
 DB 3770 GACCAAGCCAGGCTCGGAGAGCGATCCTAAAGCTGGCCGACGTGGCGGCGATCAG 3829
 QY 3126 AAGTCATGAGAGCTGCGCTCCAGGACGAGCTGAGTGGCCCTGTATACCTGAAT 3185
 DB 3830 AAGTCATGAGAGCTGCGCTCCAGGACGAGCTGAGTGGCCCTGTATACCTGAAT 3889
 QY 3186 ACCCGCAATGACCTTTTACTGCGCGACCTGCGCAAGATATACAGAGCTGCGCGA 3245
 DB 3890 ACCCGCAATGACCTTTTACTGCGCGACCTGCGCAAGATATACAGAGCTGCGCGA 3949
 QY 3246 TCCCTGATCCATTTGGGACATAGAGCGCGCAAGCCAAATTTGCAATTTGGGCGCAATTCG 3305
 DB 3950 TCCCTGATCCATTTGGGACATAGAGCGCGCAAGCCAAATTTGCAATTTGGGCGCAATTCG 4009
 QY 3306 CCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAAAGTCCCTGTGGGT 3365
 DB 4010 CCTAGTACTCTCCATTGAGCAGTAGCAGTCCCAAGCTTTGCAAAAGTCCCTGTGGGT 4069
 QY 3366 CCATTTGCTTGTTCAGAGCCACACCAACCACTTAGCATCAGCTTACTATGTCACGC 3425
 DB 4070 CCATTTGCTTGTTCAGAGCCACACCAACCACTTAGCATCAGCTTACTATGTCACGC 4129
 QY 3426 TCCCGGAGATTTCCGTTGAGCAGAGAGTGTCTTTTCTTGGAGCTGAGCATTCACAGC 3485
 DB 4130 TCCCGGAGATTTCCGTTGAGCAGAGAGTGTCTTTTCTTGGAGCTGAGCATTCACAGC 4189
 QY 3486 AACATTGAGAAAGTCCGAGAGAGATCCGAGCATGCTTCCGCGGTACAGTACAGAGCG 3545
 DB 4190 AACATTGAGAAAGTCCGAGAGAGATCCGAGCATGCTTCCGCGGTACAGTACAGAGCG 4249
 QY 3546 CTGGCGCCCAATGCTTCAATGACACTTGCAGTATCAGATCAGTACAGGCGCAACAGATTCG 3605
 DB 4250 CTGGCGCCCAATGCTTCAATGACACTTGCAGTATCAGATCAGTACAGGCGCAACAGATTCG 4309
 QY 3606 TTGTCATCTCTGTTTCAATCCAGAGCGAATCTGTCGCGCACTACATCCCAATCAAT 3665
 DB 4310 TTGTCATCTCTGTTTCAATCCAGAGCGAATCTGTCGCGCACTACATCCCAATCAAT 4369
 QY 3666 ACACCTGATGAGCAACATGAGCTGATATCTGAGCGGAGAAAGCACTCAGAAC 3725
 DB 4370 ACACCTGATGAGCAACATGAGCTGATATCTGAGCGGAGAAAGCACTCAGAAC 4429
 QY 3726 GCCAAGTCAACCACTGACGATGCCAAGGTGATCAAGGCTTCTGATAAATATGCGTGAAT 3785

DB 4430 GCCAAGTCAACCACTGACGATGCCAAGGTGATACGGTCCGATAAATATGCGTGAAT 4489
 QY 3786 GGAGAGCTGATCTGGCCAGCAACCTGATGAGAGGAAAGTGGTGGCGCTTGACG 3845
 DB 4490 GGAGAGCTGATCTGGCCAGCAACCTGATGAGAGGAAAGTGGTGGCGCTTGACG 4549
 QY 3846 CTAAACAAGATCAGCCCGTGGATGCTTCAAGCTCAGATCTTAACCTAGGGAATTCG 3905
 DB 4550 CTAAACAAGATCAGCCCGTGGATGCTTCAAGCTCAGATCTTAACCTAGGGAATTCG 4609
 QY 3906 ATCAAGGCGGAAACTGTAGCTGCCCAATAGCATCTTTAGATGATCAAGCAATGCTG 3965
 DB 4610 ATCAAGGCGGAAACTGTAGCTGCCCAATAGCATCTTTAGATGATCAAGCAATGCTG 4669
 QY 3966 CTTAACTTCTGAGGCGGAGCATAGGAGCTGGTATCGCTGGCTGACGTCAGT 4025
 DB 4670 CTTAACTTCTGAGGCGGAGCATAGGAGCTGGTATCGCTGGCTGACGTCAGT 4729
 QY 4026 AAGATTATGAGAGCAACAAATGCGTCAACCTGATGCACTTTGAGCTGATTTG 4085
 DB 4730 AAGATTATGAGAGCAACAAATGCGTCAACCTGATGCACTTTGAGCTGATTTG 4789
 QY 4086 CTGAAGTATCATGATGCTATCAACACAGCAAGAGGCTTTGGGGATATCGACTCGATG 4145
 DB 4790 CTGAAGTATCATGATGCTATCAACACAGCAAGAGGCTTTGGGGATATCGACTCGATG 4849
 QY 4146 ATACCAAGATACACCATCTTCTGATCTGTGATCTGTCAACATTTGCAATCCGTA 4205
 DB 4850 ATACCAAGATACACCATCTTCTGATCTGTGATCTGTCAACATTTGCAATCCGTA 4909
 QY 4206 ATTGCGACGGGTGATTTCCCAAAATTTGCGGCATCAAGATCTCTGTAAGTACC 4265
 DB 4910 ATTGCGACGGGTGATTTCCCAAAATTTGCGGCATCAAGATCTCTGTAAGTACC 4969
 QY 4266 GAACACCAAGTTCGAGATGATCAAGACGCCCACTTGGACATTTGTTCCCAATTCGGCG 4325
 DB 4970 GAACACCAAGTTCGAGATGATCAAGACGCCCACTTGGACATTTGTTCCCAATTCGGCG 5029
 QY 4326 CGATCGGCGGACGACACGCAATGATGATGCGAAAGCTGCGTCTTCTGATGTCAG 4385
 DB 5030 CGATCGGCGGACGACACGCAATGATGATGCGAAAGCTGCGTCTTCTGATGTCAG 5089
 QY 4386 CTGTACTTTGTTGCGGCGAGAGAAAGTCAAGCCGAAGCTGTGATGTCATTCACG 4445
 DB 5090 CTGTACTTTGTTGCGGCGAGAGAAAGTCAAGCCGAAGCTGTGATGTCATTCACG 5149
 QY 4446 AAGTTAAGTCTCTCAACGTTATCATCGAAGACACGGAACCTGATCAGTGGGGAGGA 4505
 DB 5150 AAGTTAAGTCTCTCAACGTTATCATCGAAGACACGGAACCTGATCAGTGGGGAGGA 5209
 QY 4506 AGCTTACAAAGAACTCTCCGCGGATCGTCTCATGATTTGGGAGACCTTTAATAGA 4565
 DB 5210 AGCTTACAAAGAACTCTCCGCGGATCGTCTCATGATTTGGGAGACCTTTAATAGA 5269
 QY 4566 TTCTGCTGTGCGACACAAACAAGACAGACGCGTGTCTTCCCTGCGCTGAGAG 4625
 DB 5270 TTCTGCTGTGCGACACAAACAAGACAGACGCGTGTCTTCCCTGCGCTGAGAG 5329
 QY 4626 GAAGTGAAGAGCGCGGAGCATTAATATATATTAATTAATTAATTAATTAATTAAT 4685
 DB 5330 GAAGTGAAGAGCGCGGAGCATTAATATATATTAATTAATTAATTAATTAATTAAT 5389
 QY 4686 AATTATTAACGATCATGCTGTTGTAATGCTGATGATGATATATTAATTAATTAAT 4745
 DB 5390 AATTATTAACGATCATGCTGTTGTAATGCTGATGATGATATTAATTAATTAAT 5449
 QY 4746 AATGCGCCCGAAAGAAAGTCAAGGCGCATGCGCATCTAATGTCAGAGAAATCTATT 4805
 DB 5450 AATGCGCCCGAAAGAAAGTCAAGGCGCATGCGCATCTAATGTCAGAGAAATCTATT 5509
 QY 4806 TATACATTAACGGGAATTAAGGAGTAAATCCGAAAGTACTAATTTGAGCCAAC 4865

Db 5510 TATCATTAACGGGACTTAAGCGAGTAATCCGCAAGTACTATAATTGTAGCCAC 5569
 Qy 4866 TCCGCAATCTCTACTTCTACATGAGTGTAGTGTATTTTTCACAGCTCTTAATTAAC 4925
 Db 5570 TCCGCAATCTCTACTTCTACATGAGTGTAGTGTATTTTTCACAGCTCTTAATTAAC 5569
 Qy 4826 GCTATCTCTTAAATTTGTAAGCTATATACCGCTCTTTTGTGAACTGAGAGTGTGC 4985
 Db 5530 GCTATCTCTTAAATTTGTAAGCTATATACCGCTCTTTTGTGAACTGAGAGTGTGC 5569
 Qy 4866 TTTTCTACCTTGTGTTGAGGAGCTTGTATCTTATTTGCTTCCGCTCACTAAATTC 5045
 Db 5590 TTTTCTACCTTGTGTTGAGGAGCTTGTATCTTATTTGCTTCCGCTCACTAAATTC 5749
 Qy 5046 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5105
 Db 5750 TAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 5809
 Qy 5106 AATACACAGCCCAACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 5145
 Db 5810 AATACACAGCCCAACACAGATTAATTAATTAATTAATTAATTAATTAATTAATTA 5849

RESULT 3
 AF195498
 LOCUS
 DEFINITION
 accession AF195498 5916 bp mRNA linear INV 01-NOV-2000
 version AF195498.1 GI:11066120
 keywords ras 7 (MESR7) mRNA, MESR7-3403 allele, complete cds.
 source Drosophila melanogaster (fruit fly)
 organism Drosophila melanogaster
 Eukaryote; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 5916)
 Huang, A.M., Rubin, G.M., Tsang, G., Evans-Holm, M. and Sub, C.
 Full length Drosophila melanogaster cDNA sequence
 Unpublished
 2 (bases 1 to 5916)
 Huang, A.M., Rubin, G.M., Tsang, G., Evans-Holm, M. and Sub, C.
 Direct Submission
 Submitted (18-OCT-1999) Molecular and Cell Biology, University of California at Berkeley, 545 Life Sciences Addition Bldg., Berkeley, CA 94720-3200, USA

FEATURES
 Source
 1..5916
 Location/Qualifiers
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /chromosome="3"
 /map="78C1-2"
 /clone="LD11488"
 1..5916
 /gene="MESR7"
 /allele="3403"
 771..5246
 /gene="MESR7"
 /allele="3403"
 /note="suppressor of activated Ras when overexpressed in the adult eye"
 /codon_start=1
 /product="Miseexpression suppressor of ras 7"
 /protein_id="AAG28470.1"
 /db_xref="GI:11066121"
 /translation="MAYRKPSDLDFIQMPKADMRVKVQLADLVTFSLDDTNSIVC
 TDWFLIDGLMPMTGSHFKIAQKSLAEFELIKLGSDFNAYATVLPVIVRLGDS
 RDTVREKAOILRLDMERVLVLPOLADKATSEFKRNKVRSEEFLOTIVNLEHG
 TQOLSVRVYIPVICALIDPVTNREALIOTIVELYVGVDRRLRPDRMDVYASLT
 AMLEOKRDYQKOBELLPSALKNTNGVGLDEADNIGLEPRFKRPRKRLHSAVSS
 LRKPNVDYGDAGAVTMSFSSFEVFNQNLPHAKMDMDIYKQVLVIISDKNAW
 EKRDVALKKIRALLIISYHTQPFVAVALKELSLSPVDILKEELRSQVIREACTIAY

ORIGIN
 Query Match 98.3%; Score 5055.6; DB 2; Length 5916;
 Best Local Similarity 99.3%; Pred. No. 0;
 Matches 5108; Conservative 0; Mismatches 29; Indels 5; Gaps 3;
 1
 6 CGGACAGAGCGCGTCCGAAATTTGTTCCGCGACAGACGAGCAATGACATTTGCAACG 65
 711 CACGCGAGCGGATTTGCAAAATGTTGCCGACAGACGAGCAATGACATTTGCAACG 770
 Qy 66 ATGGCTTATGGAAGCCGACGACCTTGATGCTTTATCCAGCAAAATGCCCAAGCGGAC 125
 Db 771 ATGGCTTATGGAAGCCGACGACCTTGATGCTTTATCCAGCAAAATGCCCAAGCGGAC 830
 Qy 126 ATGGCTTATGGAAGCTGACGCTGCGGAGATCTGATGATTTCTTACGACGACCAAC 185
 Db 831 ATGGCTTATGGAAGCTGACGCTGCGGAGATCTGATGATTTCTTACGACGACCAAC 890
 Qy 186 TCAATGCTGACGACGACATGGGATTCCTTATGACGCTTATGATGCCATGGTACGGGC 245
 Db 891 TCAATGCTGACGACGACATGGGATTCCTTATGACGCTTATGATGCCATGGTACGGGC 950
 Qy 246 AGCCATTTAAGATTGCAAAAAGTCCCTGAGAGCGCTTCTGAGAGCTTAATAAGCCATTG 305
 Db 951 AGCCATTTAAGATTGCAAAAAGTCCCTGAGAGCGCTTCTGAGAGCTTAATAAGCCATTG 1010
 Qy 306 GGCAGGATTTTATGATATACACAGGCTACCGTCTGACATGATGATGATGATGATGATG 365
 Db 1011 GGCAGGATTTTATGATATACACAGGCTACCGTCTGACATGATGATGATGATGATGATG 1070
 Qy 366 GACAGCAGGACACAGTCCGCGAGAAAGGCGCAATTTCTGCTGCGGACCTCATGAGACAC 425
 Db 1071 GACAGCAGGACACAGTCCGCGAGAAAGGCGCAATTTCTGCTGCGGACCTCATGAGACAC 1130
 Qy 426 AGAGTCTTCCGCGCGAGAGGCTGATGACAGAGTGGCCACTAGCTCTTCAAGCACAAG 485
 Db 1131 AGAGTCTTCCGCGCGAGAGGCTGATGACAGAGTGGCCACTAGCTCTTCAAGCACAAG 1190
 Qy 486 AACGCCAAGGTGGGAGAGGATTCCTTCAAGAGATTGGAAGCTCTTCAAGAGTACGCG 545
 Db 1191 AACGCCAAGGTGGGAGAGGATTCCTTCAAGAGATTGGAAGCTCTTCAAGAGTACGCG 1250
 Qy 546 ACCCAGCAGTTAGTGTGCGCTCTATATACCAAGTTGTGACATTTCTGAGAGATCCC 605
 Db 1251 ACCCAGCAGTTAGTGTGCGCTCTATATACCAAGTTGTGACATTTCTGAGAGATCCC 1310
 Qy 606 ACGATTAAATGAGGAGAGGCGGACATCCAAAGCTTATGGAATCTTAAAGCATGTAGGG 665
 Db 1311 ACGATTAAATGAGGAGAGGCGGACATCCAAAGCTTATGGAATCTTAAAGCATGTAGGG 1370
 Qy 666 GATCGATTGGGCGCCAGACCTCGTGCATGAGAGATGTTCTGCTCGAAATTTGGCTATG 725
 Db 1371 GATCGATTGGGCGCCAGACCTCGTGCATGAGAGATGTTCTGCTCGAAATTTGGCTATG 1430

QY 726 TTGGAGCAAAAGTTCCAGCCAGGTCAAAACAGAGGGGTCTACTGCTTACCTTACGCCCTTAA 785
DB 1431 TTGGAGCAAAAGTTCCAGCCAGGTCAAAACAGAGGGGTCTACTGCTTACCTTACGCCCTTAA 1490
QY 786 AACACGAATGGCAATGGAGTGGCTTGGACGAGCCGCAAAATTTGGGTGAGGAGCGCA 845
DB 1491 AACACGAATGGCAATGGAGTGGCTTGGACGAGCCGCAAAATTTGGGTGAGGAGCGCA 1550
QY 846 CCCACGAGATGATTAAGCGGCGCACTACACTCGGCGCTTCTGTCATCACTGCGCCCAAA 905
DB 1551 CCCACGAGATGATTAAGCGGCGCACTACACTCGGCGCTTCTGTCATCACTGCGCCCAAA 1610
QY 906 CCCAATGTGAACGATGTGACCCGGTGAATGCCGCGCGCTAAACATGGAATCTTTGCAATCT 965
DB 1611 CCCAATGTGAACGATGTGACCCGGTGAATGCCGCGCGCTAAACATGGAATCTTTGCAATCT 1670
QY 966 AGCTTTGAGTGTGCTCCGCAATTGAACATCTTCCAGCTAAGGACATGACATATCTAC 1025
DB 1671 AGCTTTGAGTGTGCTCCGCAATTGAACATCTTCCAGCTAAGGACATGACATATCTAC 1730
QY 1026 AAGCAAGTACTAGTGCATCAGTGCATTAATAAAGCAGACTGGGAGAAAGTGTGATGCT 1085
DB 1731 AAGCAAGTACTAGTGCATCAGTGCATTAATAAAGCAGACTGGGAGAAAGTGTGATGCT 1790
QY 1086 CTCGAAGATCAGGGGCACTTGTCTCATTTCTAGCTATCACTCAGCGGAGTTTGTGCT 1145
DB 1791 CTCGAAGATCAGGGGCACTTGTCTCATTTCTAGCTATCACTCAGCGGAGTTTGTGCT 1850
QY 1146 GTACAGCTAAGGAATGTGCTTAAGCTGCTGAGACATCTCAGAGAGGAATACAGATCA 1205
DB 1851 GTACAGCTAAGGAATGTGCTTAAGCTGCTGAGACATCTCAGAGAGGAATACAGATCA 1910
QY 1206 CAGGTGATCCGGAGGCGTGCATCAACATCGCTACATGCTTAAGAGCTGAGAAATATA 1265
DB 1911 CAGGTGATCCGGAGGCGTGCATCAACATCGCTACATGCTTAAGAGCTGAGAAATATA 1970
QY 1266 CTAGATGCTTCTGCTGAGCACTTTTGGACACTGTGATTAATTAATACAGAACAGCGG 1325
DB 1971 CTAGATGCTTCTGCTGAGCACTTTTGGACACTGTGATTAATTAATACAGAACAGCGG 2030
QY 1326 AAGGTCAATGCACTCCGCTTCCAAATAGCTGGAAGTATATCATTAAGTATACATGCA 1385
DB 2031 AAGGTCAATGCACTCCGCTTCCAAATAGCTGGAAGTATATCATTAAGTATACATGCA 2090
QY 1386 CCAAAAGTCTCAAGATCTACACAGACCTCTGAATCAATCAAAAGTCGAGAGCATTAAG 1445
DB 2091 CCAAAAGTCTCAAGATCTACACAGACCTCTGAATCAATCAAAAGTCGAGAGCATTAAG 2150
QY 1446 TTCACACTGTGAGCTGATGCTGCTCTTCGAGAGTGCAGACGAGGCGTTGGA 1505
DB 2151 TTCACACTGTGAGCTGATGCTGCTCTTCGAGAGTGCAGACGAGGCGTTGGA 2210
QY 1506 AGGAATGCCACCGTACTAGGAGACCTTAATAAATAATTCATGCGGAGTGCACACTGCGAT 1565
DB 2211 AGGAATGCCACCGTACTAGGAGACCTTAATAAATAATTCATGCGGAGTGCACACTGCGAT 2270
QY 1566 GCACGCGCGCATTCAGATACCGCTAATGCGCTTTCAGAGCGTCACTTTTCAGAGCTGGG 1625
DB 2271 GCACGCGCGCATTCAGATACCGCTAATGCGCTTTCAGAGCGTCACTTTTCAGAGCTGGG 2330
QY 1626 GATCAAAATATATAGAAATTTAAGCATTAAGCTGCCAGCGCATTTAGAAAGGAGAGAG 1685
DB 2331 GATCAAAATATATAGAAATTTAAGCATTAAGCTGCCAGCGCATTTAGAAAGGAGAGAG 2390
QY 1686 GCGCGCGGAG 1745
DB 2391 GCGCGCGGAG 2450
QY 1746 GTATCCCGGATGGCGGAACTTGGAACCTTGCAAAAGCCGCTAGTATGAGATCC 1805
DB 2451 GTATCCCGGATGGCGGAACTTGGAACCTTGCAAAAGCCGCTAGTATGAGATCC 2510
QY 1806 ATTTACGGGCTGACACTGCGGCTGCTCAAGAGCCAAAGTTAAGAGCGCAATATATCACTA 1865

DB 2511 ATTTACGGGCTGACACTGCGGCTGCTCAAGAGCCAAAGTTAAGCGCAATATCACTA 2570
QY 1866 TATTCAGGCAAAAGGAACTTTAGAGCCTATATTAATTCACACAGGATCGATGACAGA 1925
DB 2571 TATTCAGGCAAAAGGAACTTTAGAGCCTATATTAATTCACACAGGATCGATGACAGA 2630
QY 1926 GCAGAGCCAGTGTATCACTGCCAGACCTGCTGAAATTCGAATAGCGGTGGCACACA 1985
DB 2631 GCAGAGCCAGTGTATCACTGCCAGACCTGCTGAAATTCGAATAGCGGTGGCACACA 2690
QY 1986 GCTAACAGCCGGGATCGGTTACACACAGGCGCGGGAGAGAGATCTCAGTCA 2045
DB 2691 GCTAACAGCCGGGATCGGTTACACACAGGCGCGGGAGAGAGATCTCAGTCA 2750
QY 2046 CAGCAGAGATGAGATCCACCTCGCCAGACCAAAATGAGGATCAGTACGGTGTATT 2105
DB 2751 CAGCAGAGATGAGATCCACCTCGCCAGACCAAAATGAGGATCAGTACGGTGTATT 2810
QY 2106 GGTATTTATCTATCGTGGGCGCACTGGCGCATACCCAAAAGGCGCTCCGGAATACCCGA 2165
DB 2811 GGTATTTATCTATCGTGGGCGCACTGGCGCATACCCAAAAGGCGCTCCGGAATACCCGA 2870
QY 2166 AGCAGAGCCAGCTTAGGAGAAACGAGTCACTAGGTGAGTGTGCTTATGAAAGC 2225
DB 2871 AGCAGAGCCAGCTTAGGAGAAACGAGTCACTAGGTGAGTGTGCTTATGAAAGC 2930
QY 2226 AGTATGTACTACAGGTGCGGGGCTGAGAGTACGCGCGAGAGAGAAACCCAGTAAAG 2285
DB 2931 AGTATGTACTACAGGTGCGGGGCTGAGAGTACGCGCGAGAGAGAAACCCAGTAAAG 2990
QY 2286 CCATCGGCGCGGACGACACTGCTGGCGCAATCCGCTGAGAGAGAAATATACATTAAGCGCT 2345
DB 2991 CCATCGGCGCGGACGACACTGCTGGCGCAATCCGCTGAGAGAGAAATATACATTAAGCGCT 3050
QY 2346 GAGATGATGACAGCCAGACTATGCTTCCGGGAGCTACATGCCAGCGGCGGAATGCGG 2405
DB 3051 GAGATGATGACAGCCAGACTATGCTTCCGGGAGCTACATGCCAGCGGCGGAATGCGG 3110
QY 2406 ATGGGTAGAGCTCATGGAAGTATGATGATCTGATGACATGCACTCCGAGCGCAGTTCT 2465
DB 3111 ATGGGTAGAGCTCATGGAAGTATGATGATCTGATGACATGCACTCCGAGCGCAGTTCT 3170
QY 2466 GTGTGTTGAGAACGATCTTTGCACTCAGCTACACTAGAGTAAATTAATGAACTACTCA 2525
DB 3171 GTGTGTTGAGAACGATCTTTGCACTCAGCTACACTAGAGTAAATTAATGAACTACTCA 3230
QY 2526 CTTAGCGGAGACCAACCCGCTTGGACTGAGACAGCAGCGGCGCCATTTGACACATC 2585
DB 3231 CTTAGCGGAGACCAACCCGCTTGGACTGAGACAGCAGCGGCGCCATTTGACACATC 3290
QY 2586 GAGACGATTTATGATTCTGCGCATCGACGCAATGGTGTGTAAGAGAGATGAGCTGATC 2645
DB 3291 GAGACGATTTATGATTCTGCGCATCGACGCAATGGTGTGTAAGAGAGATGAGCTGATC 3350
QY 2646 AGCTTCACACAGTATCTGGCCGATGAGAAAGAGCTCACTCAGACGAACTTAATATGCTT 2705
DB 3351 AGCTTCACACAGTATCTGGCCGATGAGAAAGAGCTCACTCAGACGAACTTAATATGCTT 3410
QY 2706 CTGACATGTTCCGCAAGATTTATGATACACACACAGAGGTGATCTCACTGTTTTG 2765
DB 3411 CTGACATGTTCCGCAAGATTTATGATACACACACAGAGGTGATCTCACTGTTTTG 3470
QY 2766 GACACGAGTACGAACTGATCTGCTTCACTGGAAGAACTTACAGAAATGCTCTTCA 2825
DB 3471 GACACGAGTACGAACTGATCTGCTTCACTGGAAGAACTTACAGAAATGCTCTTCA 3528
QY 2826 TCTTGTGAGCGGTTGTTCAACAAATGGGACCGATCTACTCAATTCGATGACAGC 2885
DB 3529 TCTTGTGAGCGGTTGTTCAACAAATGGGACCGATCTACTCAATTCGATGACAGC 3587
QY 2886 AAGATTTGAAAGCGCTACAGGTGCTCAAGAAATTTTCCCAAGCAGTTCAGCTTAA 2945

```
Db 3588 AAGATTGGAAGCGCTAGAGGTGCTCCAGAAATATTTCCCAACGAGTTGACGTTAA 3647
QY 2946 GAACTATTTAGAAATCATATCGATTTCTACTCAAAACGCACTAACGAAGCGGCATCGCC 3005
Db 3648 GAACTATTTAGAAATCATATCGATTTCTACTCAAAACGCACTAACGAAGCGGCATCGCC 3707
QY 3006 ATTTCGCGCTTCCTTACGATCTGGGCAATATCTATGTAAGAGCGCACTTTCCGAC 3065
Db 3708 ATTTCGCGCTTCCTTACGATCTGGGCAATATCTATGTAAGAGCGCACTTTCCGAC 3767
QY 3066 GACCAAGACGAGGCGCTGAGGAGGAGCGGTCTTAAAGCTGAGCCGAGCTGGGCGAGTCAG 3125
Db 3768 GACCAAGACGAGGCGCTGAGGAGGAGCGGTCTTAAAGCTGAGCCGAGCTGGGCGAGTCAG 3827
QY 3126 AAGTCGATGAGCTGCGCTCCAGGCGAGAGCTGCTAGTGGCCCTGTATTAACCTGAAT 3185
Db 3828 AAGTCGATGAGCTGCGCTCCAGGCGAGAGCTGCTAGTGGCCCTGTATTAACCTGAAT 3887
QY 3186 ACCCGGCAATGACCTTTTATCTGGCCGACCTGCGCAAGATATATCAGGACTCTGCCGA 3245
Db 3888 ACCCGGCAATGACCTTTTATCTGGCCGACCTGCGCAAGATATATCAGGACTCTGCCGA 3947
QY 3246 TCCGATCCATTCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTG 3305
Db 3948 TCCGATCCATTCGACATGAGGCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTG 4007
QY 3306 CTTAGTACGCTTCATTTGAGCAGTACAGTCCCAAGCCTTTGCAAGTCCCTCTGCGGT 3365
Db 4008 CTTAGTACGCTTCATTTGAGCAGTACAGTCCCAAGCCTTTGCAAGTCCCTCTGCGGT 4067
QY 3366 CCATTGGCTCGCTTCAAGGCCACCACTTAAGATGACTCTTACTAGTCCAGC 3425
Db 4068 CCATTGGCTCGCTTCAAGGCCACCACTTAAGATGACTCTTACTAGTCCAGC 4127
QY 3426 TCCCGGAGCTCTCCGTGAGAGGAGGCTCTTTCTCGGAGCTGAGCAATTGAGAC 3485
Db 4128 TCCCGGAGCTCTCCGTGAGAGGAGGCTCTTTCTCGGAGCTGAGCAATTGAGAC 4187
QY 3486 AACATTGAGAAAGCTGCGAGAGAGATCCGCGACTGCTTGGGCGTCACTACAGAGCGC 3545
Db 4188 AACATTGAGAAAGCTGCGAGAGAGATCCGCGACTGCTTGGGCGTCACTACAGAGCGC 4247
QY 3546 CTGGGCGCCAAATGCTTCAATGAGACATTTGCAATGATCAGATCAGGCGCAACAGATTG 3605
Db 4248 CTGGGCGCCAAATGCTTCAATGAGACATTTGCAATGATCAGATCAGGCGCAACAGATTG 4307
QY 3606 TGTGATTCCTGCTCTTCAACTCCAAAGCGCAATGCTGGGCAACCTAACAGTCAAT 3665
Db 4308 TGTGATTCCTGCTCTTCAACTCCAAAGCGCAATGCTGGGCAACCTAACAGTCAAT 4367
QY 3666 ACACCTGAGTACGAAACATGAGGCTGATATCTGAGCGGGAAGGACCACTCAGAC 3725
Db 4368 ACACCTGAGTACGAAACATGAGGCTGATATCTGAGCGGGAAGGACCACTCAGAC 4427
QY 3726 GCCAAGTACCACTGAGATGCCAAGGTGATCAACGCTCTGATTAATATGCTGAAT 3785
Db 4428 GCCAAGTACCACTGAGATGCCAAGGTGATCAACGCTCTGATTAATATGCTGAAT 4487
QY 3786 GGAAGCTGATACGTCGCGCAAGCACTTATGAGAGCGCAAGTGTGCTGTGAG 3845
Db 4488 GGAAGCTGATACGTCGCGCAAGCACTTATGAGAGCGCAAGTGTGCTGTGAG 4547
QY 3846 CTAAACAAGATCAAGCCGCTGAGTGTGCTTCAAGAGTCACTTAACCTGGGATTTG 3905
Db 4548 CTAAACAAGATCAAGCCGCTGAGTGTGCTTCAAGAGTCACTTAACCTGGGATTTG 4607
QY 3906 ATCAAGGCGGAACTGTAGCTGCCAATAGACATTTAGATGATGATGATGATGATG 3965
Db 4608 ATCAAGGCGGAACTGTAGCTGCCAATAGACATTTAGATGATGATGATGATGATG 4667
QY 3966 CTTAACATTCGAGGCGGAGCATACGAGCTGTGATCGCTGCGCTGACGTCAGT 4025
Db 4668 CTTAACATTCGAGGCGGAGCATACGAGCTGTGATCGCTGCGCTGACGTCAGT 4727

QY 4026 AAGATTATGAGAGCAACAAATGCGTCAAACTGATGATGACTTTCTAGAGCTGATTTG 4085
Db 4728 AAGATTATGAGAGCAACAAATGCGTCAAACTGATGATGACTTTCTAGAGCTGATTTG 4787
QY 4086 CTGAAGATCATCTGATGCTATTCACACAGCAAGAGGCTTTGGCGGATATGACTGATG 4145
Db 4788 CTGAAGATCATCTGATGCTATTCACACAGCAAGAGGCTTTGGCGGATATGACTGATG 4847
QY 4146 ATACCAAGATAGACCACTCCGCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATG 4205
Db 4848 ATACCAAGATAGACCACTCCGCTGCTGATGCTGCTGATGCTGCTGATGCTGCTGATG 4907
QY 4206 ATTGCAAGGCTGATTTCCCAAAATTTGTCGCGCATCAAGATCTGCTGGAATGAC 4265
Db 4908 ATTGCAAGGCTGATTTCCCAAAATTTGTCGCGCATCAAGATCTGCTGGAATGAC 4967
QY 4266 GAACACATGCTGAGATACAGAGCCCACTTGAACATGCTGCTGCTGCTGCTGCTG 4325
Db 4968 GAACACATGCTGAGATACAGAGCCCACTTGAACATGCTGCTGCTGCTGCTGCTG 5027
QY 4326 CGATCGGCGAGACAGACGCAATCGATGCTGGAAGCTGGGCTTCTGATGCTGAC 4385
Db 5028 CGATCGGCGAGACAGACGCAATCGATGCTGGAAGCTGGGCTTCTGATGCTGAC 5087
QY 4386 CTGATCTTTGTTCTGGGCGAGAGAAAGTCAAGCCGAAAGCTGATGCTGATGCTGAC 4445
Db 5088 CTGATCTTTGTTCTGGGCGAGAGAAAGTCAAGCCGAAAGCTGATGCTGATGCTGAC 5147
QY 4446 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCAGTGGGAGAG 4505
Db 5148 AAGTTAGGCTCTCAACGTGTACATGAGAGAGCGGAACTGATCAGTGGGAGAG 5207
QY 4506 AGCTCTCAAAAGAACTCTCCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 4565
Db 5208 AGCTCTCAAAAGAACTCTCCGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTG 5267
QY 4566 TTCTGCGC--TCGTGCGCACCAACAAACAGACAGACGCGGTGCTTCCCTGCGCCTGAG 4623
Db 5268 TTCTGCGCATACAGACCAACCAACAAAGACAGACGCGGTGCTTCCCTGCGCCTGAG 5327
QY 4624 AGGAAGTGAAGAGGCGGAGCAATTAATTAATTAATTAATTAATTAATTAATTA 4683
Db 5328 AGGAAGTGAAGAGGCGGAGCAATTAATTAATTAATTAATTAATTAATTAATTA 5387
QY 4684 CTATATTATACCATCATCCGTTACTGTGTACTGTGTACTGTGTACTGTGTACTGTGT 4743
Db 5388 CTATATTATACCATCATCCGTTACTGTGTACTGTGTACTGTGTACTGTGTACTGTGT 5447
QY 4744 CAATGCGGCGCGAAAGAAAGTCAAGGCGCATGCGCATCTAATGTAGAGAGAACTA 4803
Db 5448 CAATGCGGCGCGAAAGAAAGTCAAGGCGCATGCGCATCTAATGTAGAGAGAACTA 5507
QY 4804 TTATATCATTAACGGAACCTAAAGAGTAAATCCGCAAGGTGATCTAAATTTGACCA 4863
Db 5508 TTATATCATTAACGGAACCTAAAGAGTAAATCCGCAAGGTGATCTAAATTTGACCA 5567
QY 4864 ACTCCGAATCTCATCTTCTCATCATGATGTGATGCTTTTATACAGCTCTTAATTA 4923
Db 5568 ACTCCGAATCTCATCTTCTCATCATGATGTGATGCTTTTATACAGCTCTTAATTA 5627
QY 4924 ACGCTGATCTGTTTAAATTTGTAAGCTATATACCGCTCTTATGGAACCTCAGAGCTGT 4983
Db 5628 ACGCTGATCTGTTTAAATTTGTAAGCTATATACCGCTCTTATGGAACCTCAGAGCTGT 5687
QY 4984 GCTTTTCTACCTTTGTTTCAAGGCACTTGTATCTTATTTGCTTCCGCTCACTAAT 5043
Db 5688 GCTTTTCTACCTTTGTTTCAAGGCACTTGTATCTTATTTGCTTCCGCTCACTAAT 5747
QY 5044 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5103
Db 5748 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5807
```

QY 5104 TAAATACACAGCCCAACAGATTAATCAATPAAAA 5145
DB 5808 TAAATACACAGCCCAACAGATTAATCAATPAAAA 5849

RESULT 4
AF250842 5938 bp mRNA linear INV 04-AUG-2000
LOCUS Drosophila melanogaster multiple asters (Mast) mRNA, complete cds.
DEFINITION AF250842
ACCESSION AF250842.1 GI:7650478
VERSION
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 5938)
Lemos, C.L., Sampaio, P., Maletz, H., Costa, M., Omel'yanuk, L.V.,
Liberal, V., and Sunkel, C.E.
Mast, a conserved microtubule-associated protein required for
bipolar mitotic spindle organization
EMBO J 19 (14), 3668-3682 (2000)
TITLE 2 (bases 1 to 5938)
JOURNAL
PUBMED 10899121
REFERENCE
AUTHORS Maletz, H., Lemos, C.L., Sampaio, P., and Sunkel, C.E.
TITLE Direct Submision
JOURNAL Submitted (30-MAR-2000) Genetics Molecular da Mitose, Instituto de
Biologia Molecular e Celular, Rua Campo Alegre, 823, Porto
4150-180, Portugal
FEATURES
source
1. 5938
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="3"
/map="3L: 78C1-C2"
/clone="LD11488"
/dev_stage="0-24 hours embryo"
1. 5938
/gene="Mast"
770.. 5245
/gene="Mast"
/note="Mast; microtubule-associated protein; essential for
centrosome separation and organization of mitotic
apparatus"
/codon_start=1
/product="multiple asters"
/protein_id="AF6060.1"
/db_xref="GI:7650479"

ORIGIN

Query Match 98.3%; Score 5055.6; DB 2; Length 5938;
Best Local Similarity 99.3%; Pred. No. 0; Mismatches 29; Indels 5; Gaps 3;
Matches 5108; Conservative 0;
IMRKNRHHMMHLELILKTIQCYHSEKALBDISMPRIAPSLDLSINIVN
VIATGEPFNLCALIKLEVTBHGSEITDHLIDVFNLRASBDTOSMRKAVPC
IKVQYFVIGERKVPKSLVNPISKVRLLNIVIEKORNICISGGSSSTKNSAASS"

QY 6 CGGACGAGCCGCTCCGAATTTGTTGCCGAGACGAAAGACGATGACATTGCAAG 65
DB 710 CAGCGCACCAGGATTTGAAATAGTTGCCGAGAACGAGACGATGACATTGCAAG 769
QY 66 ATGCGCTTACGAGACCCGAGGACCTGATGCTTTATCCGAAATGCCCCAGCGGAC 125
DB 770 ATGCGCTTACGAGACCCGAGGACCTGATGCTTTATCCGAAATGCCCCAGCGGAC 829
QY 126 ATGCGTGAAGAGTACAGCTCGCGGAGGATCGTGATGACATTCCTTAGGCGACGACAAAC 185
DB 830 ATGCGTGAAGAGTACAGCTCGCGGAGGATCGTGATGACATTCCTTAGGCGACGACAAAC 889
QY 186 TCAATTGCTGACGAGACATGGGATTCCTTATTTGACGAGTTTGATGACATGCTGACGAGC 245
DB 890 TCAATTGCTGACGAGACATGGGATTCCTTATTTGACGAGTTTGATGACATGCTGACGAGC 949
QY 246 AGCCACTTTAAGATTGCAAAAGTCCCTGAGAGGCGCTTCTGGAGCTAATAAGGATG 305
DB 950 AGCCACTTTAAGATTGCAAAAGTCCCTGAGAGGCGCTTCTGGAGCTAATAAGGATG 1009
QY 306 GGGACGAGTTTAAAGCATACAGGCTACCGTTCTGACCATGATGATCGGCTGGGA 365
DB 1010 GGGACGAGTTTAAAGCATACAGGCTACCGTTCTGACCATGATGATCGGCTGGGA 1069
QY 366 GACACAGAGGACACAGTCCGAGAGAGGCGCAACTTCTGCTGCGGACCTCATGAGAGC 425
DB 1070 GACACAGAGGACACAGTCCGAGAGAGGCGCAACTTCTGCTGCGGACCTCATGAGAGC 1129
QY 426 AGAGTCTTCCGCGCCGAGCGCTGATGCAAGAGCTGCGCATGAGCTTCAAGACAG 485
DB 1130 AGAGTCTTCCGCGCCGAGCGCTGATGCAAGAGCTGCGCATGAGCTTCAAGACAG 1189
QY 486 AACGCAAGAGGCGGAGAGAGTTCCTTACAGCATGTTGAAGCGCTTCATGATGACG 545
DB 1190 AACGCAAGAGGCGGAGAGAGTTCCTTACAGCATGTTGAAGCGCTTCATGATGACG 1249
QY 546 ACCGACAGCTTATGTTGCGGCTTATATACCAAGATTGTCATCTTCCGAGATCCC 605
DB 1250 ACCGACAGCTTATGTTGCGGCTTATATACCAAGATTGTCATCTTCCGAGATCCC 1309
QY 606 ACAATTATGAGGAGGAGCGGCGCATCCAAAGCTAGTGAATTTCAAGCATGTAGG 665
DB 1310 ACAATTATGAGGAGGAGCGGCGCATCCAAAGCTAGTGAATTTCAAGCATGTAGG 1369
QY 666 GATGATGCGCGCGAGACCTCGGTCGATGACATGTTCTGCTCGAAATTTGGCTATG 725
DB 1370 GATGATGCGCGCGAGACCTCGGTCGATGACATGTTCTGCTCGAAATTTGGCTATG 1429
QY 726 TTGGAGCAAAAGTTGACACAGGTCAAAAGAGGTTCACTGCTACCTTCAAGCCTTAA 785
DB 1430 TTGGAGCAAAAGTTGACACAGGTCAAAAGAGGTTCACTGCTACCTTCAAGCCTTAA 1489
QY 786 AACCAATGCGCAATGAGAGTGGCTTGAACGAGGCGGCAATATTTGGGTTGAGGAGCGA 845
DB 1490 AACCAATGCGCAATGAGAGTGGCTTGAACGAGGCGGCAATATTTGGGTTGAGGAGCGA 1549
QY 846 CCCACAGAGATGATTAAGGCGGCACTACACTGCGCGCTTCTCATCTGCGGCCAAA 905
DB 1550 CCCACAGAGATGATTAAGGCGGCACTACACTGCGCGCTTCTCATCTGCGGCCAAA 1609
QY 906 CCCAATGTAACGATGTAACGCGTATGCGCGCGCTTCAACCATGATGATCTTTCGAT 965
DB 1610 CCCAATGTAACGATGTAACGCGTATGCGCGCGCTTCAACCATGATGATCTTTCGAT 1669

Db 3827 AAGTCATGAGCTGCGCTCCAGGCCAGAGCTGCTAGTGCCCTGTATACCTGAAT 3886
 Qy 3186 ACCCGCAATAGACCTTTTACTGCGCCGACCTGCGCAAGATATACAGACTGCGCGA 3245
 Db 3887 ACCCGCAATAGACCTTTTACTGCGCGACCTGCGCAAGATATACAGACTGCGCGA 3946
 Qy 3246 TCTGTCATCCATTGGCAGATAGGGGCGCAAGCCAAAGTTGCAATTGGGGTCCCAATTG 3305
 Db 3947 TCTGTCATCCATTGGCAGATAGGGGCGCAAGCCAAAGTTGCAATTGGGGTCCCAATTG 4006
 Qy 3306 CCTAGTAGCTCTCATTTAGAGAGTAGAGTCCAAAGCTTTGCAAGCTCCCTGTGGGT 3365
 Db 4007 CTTAGTAGCTCTCATTTAGAGAGTAGAGTCCAAAGCTTTGCAAGCTCCCTGTGGGT 4066
 Qy 3366 CCAATTTGCTGCTTTCAGAGCCAGCCAGCCCACTTAGCATCAGCTTACTAGTCCAGC 3425
 Db 4067 CCAATTTGCTGCTTTCAGAGCCAGCCAGCCCACTTAGCATCAGCTTACTAGTCCAGC 4126
 Qy 3426 TCCCGGAGCTTTCCTGAGAGAGAGAGTCTTTTCTTCTGAGCTGAGCATTCAGAC 3485
 Db 4127 TCCCGGAGCTTTCCTGAGAGAGAGAGTCTTTTCTTCTGAGCTGAGCATTCAGAC 4186
 Qy 3486 AACATCAGAGAGAGTGGAGAGAGATCCGGGACCTTGGGGGTCACTCCAGAGCGG 3545
 Db 4187 AACATCAGAGAGAGTGGAGAGAGATCCGGGACCTTGGGGGTCACTCCAGAGCGG 4246
 Qy 3546 CTGGCGCCCAATGCTTCAATGACACTTGCAGTATCAGATCAGGCGCCAAAGAGATTG 3605
 Db 4247 CTGGCGCCCAATGCTTCAATGACACTTGCAGTATCAGATCAGGCGCCAAAGAGATTG 4306
 Qy 3606 TGTGCAATCCCTGCTTTTCCAACTCCAAAGCGAATCTGTCGCGCAACTACCCAGTCAAT 3665
 Db 4307 TGTGCAATCCCTGCTTTTCCAACTCCAAAGCGAATCTGTCGCGCAACTACCCAGTCAAT 4366
 Qy 3666 AACCTGAGTCAGCAACATGAGAGTGGATCTGAGAGCGGGAAGAGCACTCAGAAC 3725
 Db 4367 AACCTGAGTCAGCAACATGAGAGTGGATCTGAGAGCGGGAAGAGCACTCAGAAC 4426
 Qy 3726 GCCAAGTCACCACTGACGATGCCAAGGTGATCACTGCTCGATTAATGCTGAAAT 3785
 Db 4427 GCCAAGTCACCACTGACGATGCCAAGGTGATCACTGCTCGATTAATGCTGAAAT 4486
 Qy 3786 GGAGAGCTGATCTGCGGAGAGCACTGATGAGAGCGAAGTGGTGGCTTGAACG 3845
 Db 4487 GGAGAGCTGATCTGCGGAGAGCACTGATGAGAGCGAAGTGGTGGCTTGAACG 4546
 Qy 3846 CTACAAAGAGATGAGCCGCTGAGTGTCTGAGAGCTCACTACTACTGAGGGAATTG 3905
 Db 4547 CTACAAAGAGATGAGCCGCTGAGTGTCTGAGAGCTCACTACTACTGAGGGAATTG 4606
 Qy 3906 ATCAAGGGCGGAACTGTGAGCTGCGCAATTAAGCATTTAGATGATCATGCGATGCTG 3965
 Db 4607 ATCAAGGGCGGAACTGTGAGCTGCGCAATTAAGCATTTAGATGATCATGCGATGCTG 4666
 Qy 3966 CTTTAACCTTCTGAGGGGAGAGCATAGGAGCTGTCTCTGCGCTGCGACGCTCACT 4025
 Db 4667 CTTTAACCTTCTGAGGGGAGAGCATAGGAGCTGTCTCTGCGCTGCGACGCTCACT 4726
 Qy 4026 AAGATTATGAGAGCAACAAATGCTGACAACTGAGATGCACTTTCTAGAGCTGATTTG 4085
 Db 4727 AAGATTATGAGAGCAACAAATGCTGACAACTGAGATGCACTTTCTAGAGCTGATTTG 4786
 Qy 4086 CTGAAGATCATTCAGTGTCTATCAACAGAGAGAGGCTTTGCGGATATGACTGATG 4145
 Db 4787 CTGAAGATCATTCAGTGTCTATCAACAGAGAGAGGCTTTGCGGATATGACTGATG 4846
 Qy 4146 ATACCAAGATGATGACATCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4205
 Db 4847 ATACCAAGATGATGACATCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4906
 Qy 4206 ATTCGAGCGAGTATTCCTCAAAATTTGTCGCGCATCAAGATCTGCTGAGAGTACC 4265
 Db 4907 ATTCGAGCGAGTATTCCTCAAAATTTGTCGCGCATCAAGATCTGCTGAGAGTACC 4966

Qy 4266 GAAACACATGCTTCGAGATCAACAGCGCCCACTTGGAGATTGTGTTCCCAATTCGCG 4325
 Db 4967 GAAACACATGCTTCGAGATCAACAGCGCCCACTTGGAGATTGTGTTCCCAATTCGCG 5026
 Qy 4326 CGATCGCGGAGCAACCGCATGATGTTCCGAAAGCTGCGGTCTTTCGATTCGTCAG 4385
 Db 5027 CGATCGCGGAGCAACCGCATGATGTTCCGAAAGCTGCGGTCTTTCGATTCGTCAG 5086
 Qy 4386 CTGTACTTTGTTCTGGGCGAGAGAGTCAAGCCGAGAGCTGTCACTGCTAAATCCAG 4445
 Db 5087 CTGTACTTTGTTCTGGGCGAGAGAGTCAAGCCGAGAGCTGTCACTGCTAAATCCAG 5146
 Qy 4446 AAGTTAGGCTCTCAACGCTGTATCATGAGAGAGAGAGAGTCACTGATGAGAGAG 4505
 Db 5147 AAGTTAGGCTCTCAACGCTGTATCATGAGAGAGAGAGTCACTGATGAGAGAGAG 5206
 Qy 4506 AGCTCTCAAAAGAACTCTCCGCGGATGCTGTCTGATTTGCGGAGAGCTTTAATAGA 4565
 Db 5207 AGCTCTCAAAAGAACTCTCCGCGGATGCTGTCTGATTTGCGGAGAGCTTTAATAGA 5266
 Qy 4566 TTCTCTC--TCGCGCAGCAACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4623
 Db 5267 TTCTCTCAGTACAG 5326
 Qy 4624 AGGAGTGAAG 4683
 Db 5327 AGGAGTGAAG 5386
 Qy 4684 CTATATTAATCCGATATCCGTTACTTGTGTAACTGCTGAGAGATTAATTCAGAGAG 4743
 Db 5387 CTATATTAATCCGATATCCGTTACTTGTGTAACTGCTGAGAGATTAATTCAGAGAG 5446
 Qy 4744 CAAATGCGGCCGAGAAAGAAAGTCAAGGCGCATCGCCATCTAATGTAGAGAGAGAG 4803
 Db 5447 CAAATGCGGCCGAGAAAGAAAGTCAAGGCGCATCGCCATCTAATGTAGAGAGAGAG 5506
 Qy 4804 TTATATCATTAACGAG 4863
 Db 5507 TTATATCATTAACGAG 5566
 Qy 4864 ACTCGGCAATCTGCTCTGATCATGATGATGATGATGATGATGATGATGATGATGAT 4923
 Db 5567 ACTCGGCAATCTGCTCTGATCATGATGATGATGATGATGATGATGATGATGATGAT 5626
 Qy 4924 ACGGTATCTGTTTAATTTGTAAGCTATATACGCTTTTATGGAAGTCAAGAGTGT 4983
 Db 5627 ACGGTATCTGTTTAATTTGTAAGCTATATACGCTTTTATGGAAGTCAAGAGTGT 5686
 Qy 5044 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5103
 Db 5747 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5806
 Qy 5104 TAAATACAGAGCCAGCAAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 5145
 Db 5807 TAAATACAGAGCCAGCAAGAGATTAATTAATTAATTAATTAATTAATTAATTAAT 5848

RESULT 5
 AY069579
 LOCUS 5962 bp mRNA linear INV 17-DEC-2001
 DEFINITION Drosophila melanogaster LD31673 full length cDNA.
 ACCESSION AY069579
 VERSION AY069579.1 GI:17862493
 KEYWORDS FLI_CDNA.
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Dh 1726 AAGCAAGTACTAGTATCATGATATAAAACGACACTGGAGAAAAGTGTGATGCT 1785
Qy 1086 CTCAAGAGATCAGGGGCAATGCTCATTTCTCAGCTATCAACTCAGCCGAGTTTGTGCT 1145
Dh 1786 CTCAAGAGATCAGGGGCAATGCTCATTTCTCAGCTATCAACTCAGCCGAGTTTGTGCT 1845
Qy 1146 GTACAGCTAAAGAAATTTGCTGTTAAGCTTGTGACATCTCTCAAGAGAACTACATCA 1205
Dh 1846 GTACAGCTAAAGAAATTTGCTGTTAAGCTTGTGACATCTCTCAAGAGAACTACATCA 1905
Qy 1206 CAGGTGATCCCGGAGAGGCTGATCACTGCTGCTCATGCTTAAGACCTGAGAAATTA 1265
Dh 1906 CAGGTGATCCCGGAGAGGCTGATCACTGCTGCTCATGCTTAAGACCTGAGAAATTA 1965
Qy 1266 CTAGATGCTTCTGTGTGAGACATTTTGGAGACCTGATTAATTAATCAGAACAGCCG 1325
Dh 1966 CTAGATGCTTCTGTGTGAGACATTTTGGAGACCTGATTAATTAATCAGAACAGCCG 2025
Qy 1326 AAGGTCAATGATCCGCTTCCCAATAGCTTGAAGTATCATTAAGTATACATGCA 1385
Dh 2026 AAGGTCAATGATCCGCTTCCCAATAGCTTGAAGTATCATTAAGTATACATGCA 2085
Qy 1386 CCAAGGCTGCTCAAGATCTACACAGACACTGATATCAATCAAGTCAAGACATAAG 1445
Dh 2086 CCAAGGCTGCTCAAGATCTACACAGACACTGATATCAATCAAGTCAAGACATAAG 2145
Qy 1446 TCCACACTGTGTGAGCTGATGCTGTCTTCCAGAGAGTGGACAGACAGGCTTGGAA 1505
Dh 2146 TCCACACTGTGTGAGCTGATGCTGTCTTCCAGAGAGTGGACAGACAGGCTTGGAA 2205
Qy 1506 AGGAATGACACCGTAATAAGGAGACACTTAAATAATCATTTGGGATGACAGCTGCAT 1565
Dh 2206 AGGAATGACACCGTAATAAGGAGACACTTAAATAATCATTTGGGATGACAGCTGCAT 2265
Qy 1566 GCAGGCGCCCATTCAGATAGCCCTATGCGCTTTCAGGCGTCACTTTCCAGAGCTGGCG 1625
Dh 2266 GCAGGCGCCCATTCAGATAGCCCTATGCGCTTTCAGGCGTCACTTTCCAGAGCTGGCG 2325
Qy 1626 GATCAATATATATGAACTATAGACATAGCTGCGCCGCGCATTTAGAAAGGAGACGAG 1685
Dh 2326 GATCAATATATATGAACTATAGACATAGCTGCGCCGCGCATTTAGAAAGGAGACGAG 2385
Qy 1686 GCGCGCGGAGGAGAGAACTGTGATCTGGGACTGGGACTGCACTGAAACGACGCACT 1745
Dh 2386 GCGCGCGGAGGAGAGAACTGTGATCTGGGACTGGGACTGCACTGAAACGACGCACT 2445
Qy 1746 GTATCCCGCATTTGGCCGAAACACCTGGAAACCTGCAAAAGCCGCTATGATGATCC 1805
Dh 2446 GTATCCCGCATTTGGCCGAAACACCTGGAAACCTGCAAAAGCCGCTATGATGATCC 2505
Qy 1806 ATTTCAAGGCTGAGCACTGCGGCTGCTCAAGAGCAAGTTAGAGCCCAATATACATA 1865
Dh 2506 ATTTCAAGGCTGAGCACTGCGGCTGCTCAAGAGCAAGTTAGAGCCCAATATACATA 2565
Qy 1866 TATTTCAAGCAAAAGAACTTTTAGAGCTTATATATTTCCAAACGAGCATCATGACA 1925
Dh 2566 TATTTCAAGCAAAAGAACTTTTAGAGCTTATATATTTCCAAACGAGCATCATGACA 2625
Qy 1926 GCAGACGCAATGATCATCTGCGCCAGACCTGCGCTGAAATTCATATAGGCTGGACAC 1985
Dh 2626 GCAGACGCAATGATCATCTGCGCCAGACCTGCGCTGAAATTCATATAGGCTGGACAC 2685
Qy 1986 GCTCAACAGCGCGGATCGGTTTACACACGAGCCCGGAGAGCGGAGAGTATCTCAGTCA 2045
Dh 2686 GCTCAACAGCGCGGATCGGTTTACACACGAGCCCGGAGAGCGGAGAGTATCTCAGTCA 2745
Qy 2046 CAGCCAGAGATGAGATCACTTCCGCAAGCACAAAACTGAGGAGTCAAGTGGTGTAT 2105
Dh 2746 CAGCCAGAGATGAGATCACTTCCGCAAGCACAAAACTGAGGAGTCAAGTGGTGTAT 2805
Qy 2106 GGTATTTACTATCTGTGGGCGCACTGGCGCATATCCCAAAAAGGCTCCGGAATACCCCG 2165
Dh 2806 GGTATTTACTATCTGTGGGCGCACTGGCGCATATCCCAAAAAGGCTCCGGAATACCCCG 2865

Qy 2166 AGCAGACAGCCTCTAGAGAAAAGAGTCCAACTAGTCAAGTGTGCTTGATGAAAAGC 2225
Dh 2866 AGCAGACAGCCTCTAGAGAAAAGAGTCCAACTAGTCAAGTGTGCTTGATGAAAAGC 2925
Qy 2226 AGTATGATCTCTACAGGTGGGGGTCTCCAGCTACGCGCCGAGAGAAACAACCCGATAGA 2285
Dh 2926 AGTATGATCTCTACAGGTGGGGGTCTCCAGCTACGCGCCGAGAGAAACAACCCGATAGA 2985
Qy 2286 CCATGCGCGCGGACAGACTGTGCGCAATCCGTAAGACAGAACTACATTAAGGCTC 2345
Dh 2986 CCATGCGCGCGGACAGACTGTGCGCAATCCGTAAGACAGAACTACATTAAGGCTC 3045
Qy 2346 GAGATGATGAGACAGCCAGACTATGTTTCCGAGGACATACGCGCAGCGGATCCG 2405
Dh 3046 GAGATGATGAGACAGCCAGACTATGTTTCCGAGGACATACGCGCAGCGGATCCG 3105
Qy 2406 ATGGTATGAGAGCTCATGGAGCTGATGATCTGATGACATTCGACCGGACGATCT 2465
Dh 3106 ATGGTATGAGAGCTCATGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 3165
Qy 2466 GTGTGTTCAAGACGATCTTCCGACTCCAGCTACACTAGAGTAAATGAACTACTCA 2525
Dh 3166 GTGTGTTCAAGACGATCTTCCGACTCCAGCTACACTAGAGTAAATGAACTACTCA 3225
Qy 2526 CTTAGCGGAGCCACACCGCTTGGACTGGAGACGCGAGCGGCGCATTTGACATC 2585
Dh 3226 CTTAGCGGAGCCACACCGCTTGGACTGGAGACGCGAGCGGCGCATTTGACATC 3285
Qy 2586 GAGACGATTTATGAGTTCTGCGCATGACGCACTTGGCTGAAGAGAGTGGCTGATC 2645
Dh 3286 GAGACGATTTATGAGTTCTGCGCATGACGCACTTGGCTGAAGAGAGTGGCTGATC 3345
Qy 2646 AGCCTCAACAGTATCTGCGCCGATGAGAAAGAGACTCACTGAGAGCACTTAATGCTT 2705
Dh 3346 AGCCTCAACAGTATCTGCGCCGATGAGAAAGAGACTCACTGAGAGCACTTAATGCTT 3405
Qy 2706 CTGACATGTTCCGCAAGATGTTCAATGATATACACACCAAGGTTACTACTGTTT 2765
Dh 3406 CTGACATGTTCCGCAAGATGTTCAATGATATACACACCAAGGTTACTACTGTTT 3465
Qy 2766 GACACGCTACCGAATCGATTTCTGTTCAATGCAATGCAATGCAATGCAATGCAAT 2825
Dh 3466 GACACGCTACCGAATCGATTTCTGTTCAATGCAATGCAATGCAATGCAATGCAAT 3525
Qy 2826 TCTTGTGTCAGCGGCTTGTCAACAAATTTGGGACCGATCTATCAATGCAATGCAAT 2885
Dh 3524 TCTTGTGTCAGCGGCTTGTCAACAAATTTGGGACCGATCTATCAATGCAATGCAAT 3582
Qy 2886 AAGATTTGAGAAAGCTTACAGTGTGTCAGAAATTTTCCCAAGGTTGACGTTTAA 2945
Dh 3583 AAGATTTGAGAAAGCTTACAGTGTGTCAGAAATTTTCCCAAGGTTGACGTTTAA 3642
Qy 2946 GAATATTTAGATCAATATGAGATTTCTACAAACGCACTACCAAGACGCGATGCGC 3005
Dh 3643 GAATATTTAGATCAATATGAGATTTCTACAAACGCACTACCAAGACGCGATGCGC 3702
Qy 3006 ATTTGCGGCTTCTTACAGATCTGCGCAATATATCTGTAAGAGCAGACTTTCCAGC 3065
Dh 3703 ATTTGCGGCTTCTTACAGATCTGCGCAATATATCTGTAAGAGCAGACTTTCCAGC 3762
Qy 3066 GACCAGACGAGCCTGCGAGCGGACGCTTAAAGCTGCGCAGCTGCGCGGATCAG 3125
Dh 3763 GACCAGACGAGCCTGCGAGCGGACGCTTAAAGCTGCGCAGCTGCGCGGATCAG 3822
Qy 3126 AAGTCAATGAGCTGCGCTCCAGCGCAGAGAGTCTATGCGCTGTAATACCTAAT 3185
Dh 3823 AAGTCAATGAGCTGCGCTCCAGCGCAGAGAGTCTATGCGCTGTAATACCTAAT 3882
Qy 3186 ACCCGCAATGACCCCTTTTACTGCGCAGACTGCAAAAGTATATCAGACTCTGCGCG 3245
Dh 3883 ACCCGCAATGACCCCTTTTACTGCGCAGACTGCAAAAGTATATCAGACTCTGCGCG 3942

QY 3246 TCTGCATCATTCGACATGAGGGGCAAAAGTTCGATTCGGGTCGAATTG 3305
 DB 3943 TCTGCATCATTCGACATGAGGGGCAAAAGTTCGATTCGGGTCGAATTG 4002
 QY 3306 CCTAGTACCTCTCCATTGAGCAGTACGATCCCAAGCTTTCGAAAGTCCCTCTGGGT 3365
 DB 4003 CCTAGTACCTCTCCATTGAGCAGTACGATCCCAAGCTTTCGAAAGTCCCTCTGGGT 4062
 QY 3366 CCATTGCTCGCTTCAGAGCCCAACCACTTAGCATGACTCTAATGTCACGC 3425
 DB 4063 CCATTGCTCGCTTCAGAGCCCAACCACTTAGCATGACTCTAATGTCACGC 4122
 QY 3426 TCCCGGAGCTTCCTGAGAGAGAGCTCTTTCTCGGAGCTGAGCATTTAGAC 3485
 DB 4123 TCCCGGAGCTTCCTGAGAGAGAGCTCTTTCTCGGAGCTGAGCATTTAGAC 4182
 QY 3486 AACATTGAGAGACGTCGAGAGAGATCCGGGCTCTTCGGGGGTCAGTACAGAGCG 3545
 DB 4183 AACATTGAGAGACGTCGAGAGAGATCCGGGCTCTTCGGGGGTCAGTACAGAGCG 4242
 QY 3546 CTGGCGCCCAATGGCTTCATGACATTCGATGATCAAGATCAGGGCCAAAGGATTG 3605
 DB 4243 CTGGCGCCCAATGGCTTCATGAGACATTCGATGATCAAGATCAGGGCCAAAGGATTG 4302
 QY 3606 TGTGATCTCTGTCTTCCAACTCCAAAGCGCATGTCGGCCAACTAACCCAGTCAAT 3665
 DB 4303 TGTGATCTCTGTCTTCCAACTCCAAAGCGCATGTCGGCCAACTAACCCAGTCAAT 4362
 QY 3666 ACACCTGATGAGCAAAATGAGGCTGATTAATTCGAGCGGGGAAAGGACCACTCAGAC 3725
 DB 4363 ACACCTGATGAGCAAAATGAGGCTGATTAATTCGAGCGGGGAAAGGACCACTCAGAC 4422
 QY 3726 GCCAAGTCAACCACTGACGATGCGCAAGGTGATCAAGTCTGATTAATGCTGAAAT 3785
 DB 4423 GCCAAGTCAACCACTGACGATGCGCAAGGTGATCAAGTCTGATTAATGCTGAAAT 4482
 QY 3786 GGAAGCTGATATCTGGCCAGACCTGATGAGAGCGGAAGTGTGCTGAGCTTGAAG 3845
 DB 4483 GGAAGCTGATATCTGGCCAGACCTGATGAGAGCGGAAGTGTGCTGAGCTTGAAG 4542
 QY 3846 CTAACAAGATTCAGCCCGCTGAGTGTCTTCAAGCTGATTAATGCTGAGGATTTG 3905
 DB 4543 CTAACAAGATTCAGCCCGCTGAGTGTCTTCAAGCTGATTAATGCTGAGGATTTG 4602
 QY 3906 ATCAAGGGCGGAACTGAGCTGCGCAATAGCACTTAGATGATGATGCAATGCTG 3965
 DB 4603 ATCAAGGGCGGAACTGAGCTGCGCAATAGCACTTAGATGATGATGCAATGCTG 4662
 QY 3966 CTTAACAATCTGAGGGCGGACATACGAGCGTGTGATCGCTGGCTGACGCTCACT 4025
 DB 4663 CTTAACAATCTGAGGGCGGACATACGAGCGTGTGATCGCTGGCTGACGCTCACT 4722
 QY 4026 AAGATTATGAGAGACCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTG 4085
 DB 4723 AAGATTATGAGAGACCAAAATGCTGACAACTGAGTGCATTTCTAGAGCTGATTTG 4782
 QY 4086 CTGAAGATCATCTCAGTGTATCAACAGCAAGAGGCTTGGGGGATATGATCGATG 4145
 DB 4783 CTGAAGATCATCTCAGTGTATCAACAGCAAGAGGCTTGGGGGATATGATCGATG 4842
 QY 4146 ATACCAAGATGACCACTCTGCTGATGATGATGATGATGATGATGATGATGATG 4205
 DB 4843 ATACCAAGATGACCACTCTGCTGATGATGATGATGATGATGATGATGATGATG 4902
 QY 4206 ATTGCGACGGGTGATTTCCCAAAATTTGCGCCCATCAAGATCTGCTGGAGGTACC 4265
 DB 4903 ATTGCGACGGGTGATTTCCCAAAATTTGCGCCCATCAAGATCTGCTGGAGGTACC 4962
 QY 4266 GAACACCATGTTGAGATGACAGAGCCCACTTGATGATGATGATGATGATGATGATG 4325
 DB 4963 GAACACCATGTTGAGATGACAGAGCCCACTTGATGATGATGATGATGATGATGATG 5022
 QY 4326 CGATGGCGGACGACAGCAATGATGATGATGATGATGATGATGATGATGATGATG 4385

DB 5023 CGATGGGAGACGACGACATGATGATGATGATGATGATGATGATGATGATGATGATG 5082
 QY 4386 CTGATCTTTGTTCTGGGCGAGAGAGAGTCAAGCGGAGTGTGATGATGATGATGATG 4445
 DB 5083 CTGATCTTTGTTCTGGGCGAGAGAGAGTCAAGCGGAGTGTGATGATGATGATGATG 5142
 QY 4446 AAGTTAGGCTCTCAACGTTGATGATGAGAGAGAGGAACTGATGATGATGATGATG 4505
 DB 5143 AAGTTAGGCTCTCAACGTTGATGATGAGAGAGAGGAACTGATGATGATGATGATG 5202
 QY 4506 AGCTCTCAAAAGAACTCTCGCGGAGATGCGGATGATGATGATGATGATGATGATG 4565
 DB 5203 AGCTCTCAAAAGAACTCTCGCGGAGATGCGGATGATGATGATGATGATGATGATG 5262
 QY 4566 TTCCTGC--TCGTGCGACCAAAACAGACACAGAGCGGTTGCTTCCCTGCGCTGAGA 4623
 DB 5263 TTCCTGCATGACACACACACAAACAGACACAGAGCGGTTGCTTCCCTGCGCTGAGA 5322
 QY 4624 AGGAAGTGAAGAGCGGCGGACATTAATTAATTAATTAATTAATTAATTAATTAAT 4683
 DB 5323 AGGAAGTGAAGAGCGGCGGACATTAATTAATTAATTAATTAATTAATTAATTAAT 5382
 QY 4684 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 4743
 DB 5383 CTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5442
 QY 4744 CAAATGCGGCGCGGCGGAGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4803
 DB 5443 CAAATGCGGCGCGGAGAGGAGTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5502
 QY 4804 TTATATCATTAACGAGGAACTAAAGCGAGTAATCCGAAAGTGTACTTAATTTAGCCA 4863
 DB 5503 TTATATCATTAACGAGGAACTAAAGCGAGTAATCCGAAAGTGTACTTAATTTAGCCA 5562
 QY 4864 ACTCCGCAATCTCTCACTTCTCACTGATGATGATGATGATGATGATGATGATGATG 4923
 DB 5563 ACTCCGCAATCTCTCACTTCTCACTGATGATGATGATGATGATGATGATGATGATG 5622
 QY 4924 ACGGTATCTGTTAATTTGTAAGCTATATACCGCTCTTAATGAGAACTCAGAGCTGT 4983
 DB 5623 ACGGTATCTGTTAATTTGTAAGCTATATACCGCTCTTAATGAGAACTCAGAGCTGT 5682
 QY 4984 GCTTTTCTACCTTTGTTTCAAGCGGACTTGTATCTTAATTTGTTTCCGCTCAACTAAT 5043
 DB 5683 GCTTTTCTACCTTTGTTTCAAGCGGACTTGTATCTTAATTTGTTTCCGCTCAACTAAT 5742
 QY 5044 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5103
 DB 5743 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5802
 QY 5104 TAAATACAAAGCCACAAAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 5145
 DB 5803 TAAATACAAAGCCACAAAGGATTAATTAATTAATTAATTAATTAATTAATTAAT 5844

RESULT 6
 AC014071/c 87197 bp DNA linear HTG 16-NOV-1999
 LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 DEFINITION AC014071
 ACCESSION AC014071.1 GI:6437264
 VERSION HTG: HTGS PHASE2
 KEYWORDS Drosophila melanogaster (fruit fly)
 SOURCE Drosophila melanogaster
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 87197)
 AUTHORS Adams, M. and Venter, J.C.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA

COMMENT

This sequence was identified as CDM:10214235 by the submitter.
For further information on this sequence e-mail to fly@celera.com.

- * NOTE: This is a 'working draft' sequence.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

FEATURES

source

1. 87197
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN

```
Query Match      98.2%; Score 5051.8; DB 14; Length 87197;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

OY 29 GTTGGCCGAGAACGAGACGATAGACATTGCAACGATGCGCTTTCGGAAGCCGACGCA 88
Db 65836 GTTGGCCGAGAACGAGACGATAGACATTGCAACGATGCGCTTTCGGAAGCCGACGCA 65777
OY 89 CCTGATGAGCTTATCCAGCAAAATGCCAAGGCGGACATGCGTGTGAAGGTACAGCTCGC 148
Db 65776 CCTGATGAGCTTATCCAGCAAAATGCCAAGGCGGACATGCGTGTGAAGGTACAGCTCGC 65717
OY 149 CGAGATCTGTGATCTCTTAGCGAGACACAAACTCAATTGTCTGACGCGACATGGG 208
Db 65716 CGAGATCTGTGATCTCTTAGCGAGACACAAACTCAATTGTCTGACGCGACATGGG 65657
OY 209 ATTCTTATTTACGCGTTTGATGTCATGGCTGACGCGGACCCCTTTAGATTGACAAAA 268
Db 65656 ATTCTTATTTACGCGTTTGATGTCATGGCTGACGCGGACCCCTTTAGATTGACAAAA 65597
OY 269 GTCCCTGAGAGGCGTCTCGAGACTAATAAGCGATTGGGACGCAATTTTATGATACAC 328
Db 65596 GTCCCTGAGAGGCGTCTCGAGACTAATAAGCGATTGGGACGCAATTTTATGATACAC 65537
OY 329 GGCCTACCGTTCTGCACATGTGATCGATCGCTGGGAGACAGCAGGAGACACAGTCGCGCA 388
Db 65536 GGCCTACCGTTCTGCACATGTGATCGATCGCTGGGAGAGACAGCAGGAGACACAGTCGCGCA 65477
OY 389 GAAAGCGCACTTTCGCTGCGGACCTCTAGAGACACAGATGCTTCCGCCACAGCGCT 448
Db 65476 GAAAGCGCACTTTCGCTGCGGACCTCTAGAGACACAGATGCTTCCGCCACAGCGCT 65417
OY 449 GATCGACAAGCTGCGACATGCTGCTTCAAGCACAAGAACGCCAAGGTGCGGAGAGATT 508
Db 65416 GATCGACAAGCTGCGACATGCTGCTTCAAGCACAAGAACGCCAAGGTGCGGAGAGATT 65357
OY 509 CTTTCAGACGATTTGGAACGCTCTTCATGAGTACGCGACCCAGCAGCTTATGTTCCGCT 568
Db 65356 CTTTCAGACGATTTGGAACGCTCTTCATGAGTACGCGACCCAGCAGCTTATGTTCCGCT 65297
OY 569 CTATATACCAAGGTTTGTGCACTTCTGGAAGATCCCAAGTTATTTGAGGAGAGCGGC 628
Db 65296 CTATATACCAAGGTTTGTGCACTTCTGGAAGATCCCAAGTTATTTGAGGAGAGCGGC 65237
OY 629 CATCAAAAGCTAGTGAAGAACTACAGAGATGTAGGGATCGATGCGGCCGACAGCTCCG 688
Db 65236 CATCAAAAGCTAGTGAAGAACTACAGAGATGTAGGGATCGATGCGGCCGACAGCTCCG 65177
OY 689 TCGCATGAGCAGATGTTCTCGCTCGAAATTTGGCTATGTTGAGCAAAAGTTTCAGCAGAT 748
Db 65176 TCGCATGAGCAGATGTTCTCGCTCGAAATTTGGCTATGTTGAGCAAAAGTTTCAGCAGAT 65117
OY 749 CAAAGAGAGGCTCTACTGCTACTCTTCAAGCCCTTAAAAACAGAAATGGGAATGGAGTGG 808
Db 65116 CAAAGAGAGGCTCTACTGCTACTCTTCAAGCCCTTAAAAACAGAAATGGGAATGGAGTGG 65057
OY 809 CTTGACGAGGCGCAGCAATATTTGGGTTGAGGAGGAGCCACCAAGATGATTTAAGCGGCC 868
Db 65056 CTTGACGAGGCGCAGCAATATTTGGGTTGAGGAGGAGCCACCAAGATGATTTAAGCGGCC 64997
```

```
OY 869 ACTACATCTGGCCCTTTTCGTCATCTACCTGCGCCCAAAACCCAAATGTGAACGATGACCGG 928
Db 64996 ACTACATCTGGCCCTTTTCGTCATCTACCTGCGCCCAAAACCCAAATGTGAACGATGACCGG 64937
OY 929 TGAATGCGGCGCGGTAAACATGGAATCTTTGGAATCTAGCTTTGAGGTGTCGCCCAATT 988
Db 64936 TGAATGCGGCGCGGTAAACATGGAATCTTTGGAATCTAGCTTTGAGGTGTCGCCCAATT 64877
OY 989 GAACATCTTCCACGCTAAGGACATGACGATATCTACAAAGCACTAGTATGATCATCAG 1048
Db 64876 GAACATCTTCCACGCTAAGGACATGACGATATCTACAAAGCACTAGTATGATCATCAG 64817
OY 1049 TGATAAAAAGCGACAGCTGGGAGAAAGCTGTGATGCTCTCAAGAAATCAGGGATTCCT 1108
Db 64816 TGATAAAAAGCGACAGCTGGGAGAAAGCTGTGATGCTCTCAAGAAATCAGGGATTCCT 64757
OY 1109 CATCTGAGCTATGACCTGACGCGGAGTTTGTGCTGTACAGTAAAGAAATTTGTCGTT 1168
Db 64756 CATCTGAGCTATGACCTGACGCGGAGTTTGTGCTGTACAGTAAAGAAATTTGTCGTT 64697
OY 1169 AAGCTTGTGAGACATCTCTCAAGAGAACTACATCAAGGTATCCGAGGCGTGCAT 1228
Db 64696 AAGCTTGTGAGACATCTCTCAAGAGAACTACATCAAGGTATCCGAGGCGTGCAT 64637
OY 1229 CACCATGCTCTACATGCTCTAAGAGCTGAGAAATTAATCTAGATGCTTGTGAGAGAT 1288
Db 64636 CACCATGCTCTACATGCTCTAAGAGCTGAGAAATTAATCTAGATGCTTGTGAGAGAT 64577
OY 1289 TTTGAGACACCTGATTAATTTAATTAATACAGAAACAGGCGCAAGGTCTTCATCCGCTTCAC 1348
Db 64576 TTTGAGACACCTGATTAATTTAATTAATTAATACAGAAACAGGCGCAAGGTCTTCATCCGCTTCAC 64517
OY 1349 AATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTACAC 1408
Db 64516 AATAGCTCTGAAGTATATCATTAAGTATACATGACACCAAGCTGCTCAAGATCTACAC 64457
OY 1409 AGACATCTGAATTCATCAAAAGTGAAGACATTAAGGTCCACATGCTGATGATGATGCT 1468
Db 64456 AGACATCTGAATTCATCAAAAGTGAAGACATTAAGGTCCACATGCTGATGATGATGCT 64397
OY 1469 GCTGCTCTTGAAGAGTGGGAGACGAAAGCGTGTGAAGAAAGATCCACCTGATTAAGGGA 1528
Db 64396 GCTGCTCTTGAAGAGTGGGAGACGAAAGCGTGTGAAGAAAGATCCACCTGATTAAGGGA 64337
OY 1529 CACCTTAAAAAATTCATTTGGCGATGACAGCTGCGATGACCGCCCATTCACATACGC 1588
Db 64336 CACCTTAAAAAATTCATTTGGCGATGACAGCTGCGATGACCGCCCATTCACATACGC 64277
OY 1589 CTATTTGGGCTTTCAAGGCTCACTTTCCAGAGCTGGCGGATCAATATTTATGGAACATTA 1648
Db 64276 CTATTTGGGCTTTCAAGGCTCACTTTCCAGAGCTGGCGGATCAATATTTATGGAACATTA 64217
OY 1649 CATAGCTGCGCAGCGCATTTAGAAAGGAAAGGAGCGGCGGAGAGAGAGAACTGG 1708
Db 64216 CATAGCTGCGCAGCGCATTTAGAAAGGAAAGGAGCGGCGGAGAGAGAGAACTGG 64157
OY 1709 TACTGGAGCTGGGACTGCACTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1768
Db 64156 TACTGGAGCTGGGACTGCACTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 64097
OY 1769 TGGAAACCTGGAAGAGCGCGCTAGTATGAGATTCATTTCAAGGGGAGACATGCGCGC 1828
Db 64096 TGGAAACCTGGAAGAGCGCGCTAGTATGAGATTCATTTCAAGGGGAGACATGCGCGC 64037
OY 1829 TGTCTCAAGAGCGCAAGTTAGAGCGCAATATACATATTTCCAGGCAAGAGAAACCTTT 1888
Db 64036 TGTCTCAAGAGCGCAAGTTAGAGCGCAATATATACATATTTCCAGGCAAGAGAAACCTTT 63977
OY 1889 AGGACCTAATATTTCCAAACGAGCATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 1948
Db 63976 AGGACCTAATATTTCCAAACGAGCATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAG 63917
OY 1949 CAGACCTGCGCTGAATTCATATAGCGGTGGCAACAGCTACAAAGCGCGGAGATGGTTAC 2008
```

Db 63916 CAGACCTCGCTGAAATTCATTAAGCGTGGACAACACACTCAACACGCCGGAGTGGTTAC 63857
 Qy 2009 ACCACGGCCCCGGGAGCGGAGAGTATCTCAGTCAAGCAGGATCGAGATCCACCTC 2068
 Db 63856 ACCACGGCCCCGGGAGCGGAGAGTATCTCAGTCAAGCAGGATCGAGATCCACCTC 63797
 Qy 2069 GCCAAGCACAATACTGAGGGATCAGTACGGTGGTATTTGTAATTAATCTCGTGGGGCCAC 2128
 Db 63796 GCCAAGCACAATACTGAGGGATCAGTACGGTGGTATTTGTAATTAATCTCGTGGGGCCAC 63737
 Qy 2129 TGGGGCCATACCCAAAAGGCTCCGGAATACCCGGAAGCAGCCAGCTTTAGGAAAC 2188
 Db 63736 TGGGGCCATACCCAAAAGGCTCCGGAATACCCGGAAGCAGCCAGCTTTAGGAAAC 63677
 Qy 2189 GAGTCCAACTAGTACAGTGGTGGCTTGTATGAAAAGCAGTATGTAAGTCAAGGTGGCGG 2248
 Db 63676 GAGTCCAACTAGTACAGTGGTGGCTTGTATGAAAAGCAGTATGTAAGTCAAGGTGGCGG 63617
 Qy 2249 GTCTCGACGTACGCCCGAGAGGAACAACCCAGTAAGACATCGCGCGCGGACGACATGCT 2308
 Db 63616 GTCTCGACGTACGCCCGAGAGGAACAACCCAGTAAGACATCGCGCGCGGACGACATGCT 63557
 Qy 2309 GGGCCATATCCGTGTAAGGAGAACTACTATGCGGTGGAGATGATGACAGCCAGACTA 2368
 Db 63556 GGGCCATATCCGTGTAAGGAGAACTACTATGCGGTGGAGATGATGACAGCCAGACTA 63497
 Qy 2369 TGTTCGGGGGACTACATGCGACGCGGGGGAATGGGTGGAGGCTCATGGGACG 2428
 Db 63496 TGTTCGGGGGACTACATGCGACGCGGGGGAATGGGTGGAGGCTCATGGGACG 63437
 Qy 2429 TGATGATCTGTGATGATCATCGACTCCGAGGCCAGTCTGTGTGTTCAAGACGATCTTTGCA 2488
 Db 63436 TGATGATCTGTGATGATCATCGACTCCGAGGCCAGTCTGTGTGTTCAAGACGATCTTTGCA 63377
 Qy 2489 CTCCAGCTCACTAGAGTAAATAATGAACTACTACTTACGGGAGCCACACCCGCTT 2548
 Db 63376 CTCCAGCTCACTAGAGTAAATAATGAACTACTACTTACGGGAGCCACACCCGCTT 63317
 Qy 2549 GGAATGAGACGAGAGCGGGGCCATTTGACGACATGAGACGATTAATTCAGTCTGCGC 2608
 Db 63316 GGAATGAGACGAGAGCGGGGCCATTTGACGACATGAGACGATTAATTCAGTCTGCGC 63257
 Qy 2609 ATGACGCATTTGCTGTAAGGAGATGAGCTGATGACCTCAACAGTATCTGGCGA 2668
 Db 63256 ATGACGCATTTGCTGTAAGGAGATGAGCTGATGACCTCAACAGTATCTGGCGA 63197
 Qy 2669 TGGAAAGAGCTCACTCAGCAGCACTTAATGCGTTTGGACATGTTCCGCAAGATGTT 2728
 Db 63196 TGGAAAGAGCTCACTCAGCAGCACTTAATGCGTTTGGACATGTTCCGCAAGATGTT 63137
 Qy 2729 CATGATACACACCAAGGATCACTGTTTGGACAGGTCACCGAATGATTC 2788
 Db 63136 CATGATACACACCAAGGATCACTGTTTGGACAGGTCACCGAATGATTC 63077
 Qy 2789 GGTTCATGCGAAGAACTTCACGAAATGCGCTTCTCACTCTGTTGAAGCGGTTGTTCAA 2848
 Db 63076 GGTTCATGCGAAGAACTTCACGAAATGCGCTTCTCACTCTGTTGAAGCGGTTGTTCAA 63020
 Qy 2849 CAATTTGGGACCGATCTCAATTTGATGACAGCAAGATTTGGAAGCGCTACAGGT 2908
 Db 63019 CAATTTGGGACCGATCTCAATTTGATGACAGCAAGATTTGGAAGCGCTACAGGT 62960
 Qy 2909 GGTCACGAATATTTCCCAAGGAGTTGACGCTTAAGAACTATTTGAATCATATGGA 2968
 Db 62959 GGTCACGAATATTTCCCAAGGAGTTGACGCTTAAGAACTATTTGAATCATATGGA 62900
 Qy 2969 TTCTACTCAAAAGCCTCACTACAGAGAGCGCGATCGCATTCGCTTCTTCAAGATCT 3028
 Db 62899 TTCTACTCAAAAGCCTCACTACAGAGAGCGCGATCGCATTCGCTTCTTCAAGATCT 62840
 Qy 3029 GGGCAATACATCTGTAAAGAGGACTTTCCAGGACCAAGGCCGCTCGAGCG 3088
 Db 62839 GGGCAATACATCTGTAAAGAGGACTTTCCAGGACCAAGGCCGCTCGAGCG 62780
 Qy 3089 GAGGCTCTTAAAGCTGGCCAGCTGGCGCGGATCAAGAACTGATGAGTGGCTCCCA 3148
 Db 62779 GAGGCTCTTAAAGCTGGCCAGCTGGCGCGGATCAAGAACTGATGAGTGGCTCCCA 62720
 Qy 3149 GGGCAGAGCTGCTTACTGAGGCTCTTAACTGAAATACCCCGCAATGACCTTTTACT 3208
 Db 62719 GGGCAGAGCTGCTTACTGAGGCTCTTAACTGAAATACCCCGCAATGACCTTTTACT 62660
 Qy 3209 GGGCAGCTGCTCAAAAGGATATCAAGAGCTGAGCCGATTCCTGATTCATTCGACATGAG 3268
 Db 62659 GGGCAGCTGCTCAAAAGGATATCAAGAGCTGAGCCGATTCCTGATTCATTCGACATGAG 62600
 Qy 3269 GCGCAAGCCTCAAAAGTTCGAGTCCGAGTCCGATTCGATTCGACATGAG 3328
 Db 62599 GCGCAAGCCTCAAAAGTTCGAGTCCGAGTCCGATTCGATTCGACATGAG 62540
 Qy 3329 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGTCAATTTGCTCGCTTCAGAGCCA 3388
 Db 62539 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGTCAATTTGCTCGCTTCAGAGCCA 62480
 Qy 3389 CCAACACCACTTAGATCAGCTCTACTAGTCCAGCGCTCCGCGAGCTTCGCGAGCA 3448
 Db 62479 CCAACACCACTTAGATCAGCTCTACTAGTCCAGCGCTCCGCGAGCTTCGCGAGCA 62420
 Qy 3449 GGAAGTCTCTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAGCTCGAGGA 3508
 Db 62419 GGAAGTCTCTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAGCTCGAGGA 62360
 Qy 3509 GATCCGCACTGCTTGGCGGTGATGACAGAGCGGCTGGCGCCATGCTTCAATGG 3568
 Db 62359 GATCCGCACTGCTTGGCGGTGATGACAGAGCGGCTGGCGCCATGCTTCAATGG 62300
 Qy 3569 ACACTTGAGTATCAAGATAGGGCCAAAGATGCTGTCATTCCTGCTTCAATCTC 3628
 Db 62299 ACACTTGAGTATCAAGATAGGGCCAAAGATGCTGTCATTCCTGCTTCAATCTC 62240
 Qy 3629 CAAGAGCAATCGTCGGCCAACTACCCAGTCAAAATCACTGATCAGCAACATGAG 3688
 Db 62239 CAAGAGCAATCGTCGGCCAACTACCCAGTCAAAATCACTGATCAGCAACATGAG 62180
 Qy 3689 GCTGATATCTGAGCGGAAAGAACCACTCAAGAGCCAGTACCACTGATGCTC 3748
 Db 62179 GCTGATATCTGAGCGGAAAGAACCACTCAAGAGCCAGTACCACTGATGCTC 62120
 Qy 3749 CAAGGATACAGGCTCCGATAAATATGCTGTAATGAGAGCTGATCTGCGCAGCA 3808
 Db 62119 CAAGGATACAGGCTCCGATAAATATGCTGTAATGAGAGCTGATCTGCGCAGCA 62060
 Qy 3809 CTTGATGAGAGCGAGTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 3868
 Db 62059 CTTGATGAGAGCGAGTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 62000
 Qy 3869 GTTGCTTCAAGCTCACTTAACCTGGGATTTGATCAAGGGCGAAATCTGTAGCT 3928
 Db 61999 GTTGCTTCAAGCTCACTTAACCTGGGATTTGATCAAGGGCGAAATCTGTAGCT 61940
 Qy 3929 GGGCAATAGCACTTTAGATGATGATGCAATGCAAGCTTGAAGCTTGAAGCTTGA 3988
 Db 61939 GGGCAATAGCACTTTAGATGATGATGCAATGCAAGCTTGAAGCTTGAAGCTTGA 61880
 Qy 3989 TACGAGAGTGTATGCTGCTGCTGCAAGCTGCTCAAGTAAAGATTTGAGAGCAAAAT 4048
 Db 61879 TACGAGAGTGTATGCTGCTGCTGCAAGCTGCTCAAGTAAAGATTTGAGAGCAAAAT 61820
 Qy 4049 GCGTCAACATGATGATCTTTAGAGCTGATTTGCTGAGATCATTCAGTGTATCA 4108
 Db 61819 GCGTCAACATGATGATCTTTAGAGCTGATTTGCTGAGATCATTCAGTGTATCA 61760
 Qy 4109 ACAAGCAAGAGAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4168
 Db 61759 ACAAGCAAGAGAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 61700

Db 62839 GGGCAATACATCTGCAAAAGCAGCGACTTTCCAGGACCAAGCCAGGCTCGAGCG 62780
 Qy 3089 GAGGCTCTTAAAGCTGGCCAGCTGGCGCGGATCAAGAACTGATGAGTGGCTCCCA 3148
 Db 62779 GAGGCTCTTAAAGCTGGCCAGCTGGCGCGGATCAAGAACTGATGAGTGGCTCCCA 62720
 Qy 3149 GGGCAGAGCTGCTTACTGAGGCTCTTAACTGAAATACCCCGCAATGACCTTTTACT 3208
 Db 62719 GGGCAGAGCTGCTTACTGAGGCTCTTAACTGAAATACCCCGCAATGACCTTTTACT 62660
 Qy 3209 GGGCAGCTGCTCAAAAGGATATCAAGAGCTGAGCCGATTCCTGATTCATTCGACATGAG 3268
 Db 62659 GGGCAGCTGCTCAAAAGGATATCAAGAGCTGAGCCGATTCCTGATTCATTCGACATGAG 62600
 Qy 3269 GCGCAAGCCTCAAAAGTTCGAGTCCGAGTCCGATTCGATTCGACATGAG 3328
 Db 62599 GCGCAAGCCTCAAAAGTTCGAGTCCGAGTCCGATTCGATTCGACATGAG 62540
 Qy 3329 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGTCAATTTGCTCGCTTCAGAGCCA 3388
 Db 62539 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGGGTCAATTTGCTCGCTTCAGAGCCA 62480
 Qy 3389 CCAACACCACTTAGATCAGCTCTACTAGTCCAGCGCTCCGCGAGCTTCGCGAGCA 3448
 Db 62479 CCAACACCACTTAGATCAGCTCTACTAGTCCAGCGCTCCGCGAGCTTCGCGAGCA 62420
 Qy 3449 GGAAGTCTCTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAGCTCGAGGA 3508
 Db 62419 GGAAGTCTCTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAGCTCGAGGA 62360
 Qy 3509 GATCCGCACTGCTTGGCGGTGATGACAGAGCGGCTGGCGCCATGCTTCAATGG 3568
 Db 62359 GATCCGCACTGCTTGGCGGTGATGACAGAGCGGCTGGCGCCATGCTTCAATGG 62300
 Qy 3569 ACACTTGAGTATCAAGATAGGGCCAAAGATGCTGTCATTCCTGCTTCAATCTC 3628
 Db 62299 ACACTTGAGTATCAAGATAGGGCCAAAGATGCTGTCATTCCTGCTTCAATCTC 62240
 Qy 3629 CAAGAGCAATCGTCGGCCAACTACCCAGTCAAAATCACTGATCAGCAACATGAG 3688
 Db 62239 CAAGAGCAATCGTCGGCCAACTACCCAGTCAAAATCACTGATCAGCAACATGAG 62180
 Qy 3689 GCTGATATCTGAGCGGAAAGAACCACTCAAGAGCCAGTACCACTGATGCTC 3748
 Db 62179 GCTGATATCTGAGCGGAAAGAACCACTCAAGAGCCAGTACCACTGATGCTC 62120
 Qy 3749 CAAGGATACAGGCTCCGATAAATATGCTGTAATGAGAGCTGATCTGCGCAGCA 3808
 Db 62119 CAAGGATACAGGCTCCGATAAATATGCTGTAATGAGAGCTGATCTGCGCAGCA 62060
 Qy 3809 CTTGATGAGAGCGAGTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 3868
 Db 62059 CTTGATGAGAGCGAGTGGTGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 62000
 Qy 3869 GTTGCTTCAAGCTCACTTAACCTGGGATTTGATCAAGGGCGAAATCTGTAGCT 3928
 Db 61999 GTTGCTTCAAGCTCACTTAACCTGGGATTTGATCAAGGGCGAAATCTGTAGCT 61940
 Qy 3929 GGGCAATAGCACTTTAGATGATGATGCAATGCAAGCTTGAAGCTTGAAGCTTGA 3988
 Db 61939 GGGCAATAGCACTTTAGATGATGATGCAATGCAAGCTTGAAGCTTGAAGCTTGA 61880
 Qy 3989 TACGAGAGTGTATGCTGCTGCTGCAAGCTGCTCAAGTAAAGATTTGAGAGCAAAAT 4048
 Db 61879 TACGAGAGTGTATGCTGCTGCTGCAAGCTGCTCAAGTAAAGATTTGAGAGCAAAAT 61820
 Qy 4049 GCGTCAACATGATGATCTTTAGAGCTGATTTGCTGAGATCATTCAGTGTATCA 4108
 Db 61819 GCGTCAACATGATGATCTTTAGAGCTGATTTGCTGAGATCATTCAGTGTATCA 61760
 Qy 4109 ACAAGCAAGAGAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 4168
 Db 61759 ACAAGCAAGAGAGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGAT 61700

QY 269 GTCCCTGAGGGGCTTCTCGAGCTAATAAGGATTTGGGACAGATTTTATGATACAC 328
Db 59511 GTCCCTGAGGGGCTTCTCGAGCTAATAAGGATTTGGGACAGATTTTATGATACAC 59532
QY 329 GGCTACCGTCTTGCCACATGTATGATGTGGGAGACAGAGGACACAGTCCGGA 388
Db 59531 GGCTACCGTCTTGCCACATGTATGATGTGGGAGACAGAGGACACAGTCCGGA 59472
QY 389 GAAGCGGACGACTTGTGCTGGCGGACCTCATGAGGACAGAGTCTTCCGCCAGGCGCT 448
Db 59471 GAAGCGGACGACTTGTGCTGGCGGACCTCATGAGGACAGAGTCTTCCGCCAGGCGCT 59412
QY 449 GATCGACAAGCTGGCCACTAGTGTCTTCAAGCACAAGACGCGCAAGTGGCGGAGAGT 508
Db 59411 GATCGACAAGCTGGCCACTAGTGTCTTCAAGCACAAGACGCGCAAGTGGCGGAGAGT 59352
QY 509 CCTTCAGACGATTTGAAAGCTCTTCAATGATGACGGACCCAGCAGCTTATGTTCGCT 568
Db 59351 CCTTCAGACGATTTGAAAGCTCTTCAATGATGACGGACCCAGCAGCTTATGTTCGCT 59292
QY 569 CTATATACCAACAGTTTGTGACCTTCTGGAGATCCCAAGTTATGTGAGGAGGAGGCG 628
Db 59291 CTATATACCAACAGTTTGTGACCTTCTGGAGATCCCAAGTTATGTGAGGAGGAGGCG 59232
QY 629 CATCCAAACGCTAGTGAATAATCTACAGCATGTAGGGATCGATTGCCCGACAGCTTCG 688
Db 59231 CATCCAAACGCTAGTGAATAATCTACAGCATGTAGGGATCGATTGCCCGACAGCTTCG 59172
QY 689 TCGCATGACGATGTCTCTGCTCGAATTTGGCTATGTTGAGCAAAAAGTTTCAGCAGT 748
Db 59171 TCGCATGACGATGTCTCTGCTCGAATTTGGCTATGTTGAGCAAAAAGTTTCAGCAGT 59112
QY 749 CAAACAGAGGGGTACTGCTACTCTTCAAGCCCTTAAAAACAGATTTGCAATGAGTGG 808
Db 59111 CAAACAGAGGGGTACTGCTACTCTTCAAGCCCTTAAAAACAGATTTGCAATGAGTGG 59052
QY 809 CTTGACGAGGCGGCAATATTGGGTGAGGAGGACCCACAGATGATTAAAGCGGCG 868
Db 59051 CTTGACGAGGCGGCAATATTGGGTGAGGAGGACCCACAGATGATTAAAGCGGCG 58992
QY 869 ACTACACTGCGCGCTTTTGTGTATCATCTGCGCCCAAAACCCATGTGAACGATGTACCG 928
Db 58991 ACTACACTGCGCGCTTTTGTGTATCATCTGCGCCCAAAACCCATGTGAACGATGTACCG 58932
QY 929 TGATGCGGCGCGGCAACATGGAATCTTGAATCTAGCTTGAAGTGGTCCGCAAT 988
Db 58931 TGATGCGGCGCGGCAACATGGAATCTTGAATCTAGCTTGAAGTGGTCCGCAAT 58872
QY 989 GAACATCTTCCACGCTAAGACATGAGCATATCTACAAGCAAGTACTAGTATCATGAG 1048
Db 58871 GAACATCTTCCACGCTAAGACATGAGCATATCTACAAGCAAGTACTAGTATCATGAG 58812
QY 1049 TGATTAACACGCACTGGAGAAAGCTGTGATGCTCTCAAGAGATCAGGCGATTGCT 1108
Db 58811 TGATTAACACGCACTGGAGAAAGCTGTGATGCTCTCAAGAGATCAGGCGATTGCT 58752
QY 1109 CATCTCGACTATCACTACAGCGGAGTTGTGCTGTGACGATCAAGAAAGAAATGTGCT 1168
Db 58751 CATCTCGACTATCACTACAGCGGAGTTGTGCTGTGACGATCAAGAAAGAAATGTGCT 58652
QY 1169 AAGCTTCGTGACATCTCAAGAGAACTACGATCAGAGTATCCGCGAGCGGTGAT 1228
Db 58651 AAGCTTCGTGACATCTCAAGAGAACTACGATCAGAGTATCCGCGAGCGGTGAT 58612
QY 1229 CACCATTCGCTTCAATGTCTAAGACGCTGAGAAATACTAATGCTTGTGTGAGCAT 1288
Db 58611 CACCATTCGCTTCAATGTCTAAGACGCTGAGAAATACTAATGCTTGTGTGAGCAT 58572
QY 1289 TTTGGAGACCTGATTAATTTATACAGAACGCGGAAAGTCAATTGCATCCGTTCCAC 1348
Db 58571 TTTGGAGACCTGATTAATTTATACAGAACGCGGAAAGTCAATTGCATCCGTTCCAC 58512

QY 1349 AATAGCTCTGAAGTATATCATTAAGTATACATGCAACCAAGCTGCTCAGATCTTAC 1408
Db 58511 AATAGCTCTGAAGTATATCATTAAGTATACATGCAACCAAGCTGCTCAGATCTTAC 58452
QY 1409 AGACACTGTGAATCAATCAAAAGTGAAGACATTAAGTCCACATGTGTGAGCTGATGAT 1468
Db 58451 AGACACTGTGAATCAATCAAAAGTGAAGACATTAAGTCCACATGTGTGAGCTGATGAT 58392
QY 1469 GCTGCTCTTCAAGAGAGTGGCAGACGAAGCGCTTGGAAAAGAAATCCACTTAAGGGA 1528
Db 58391 GCTGCTCTTCAAGAGAGTGGCAGACGAAGCGCTTGGAAAAGAAATCCACTTAAGGGA 58332
QY 1529 CACCTTAATAAATTCATTTGGCGATGACAGTCCGATGCAAGCCGCTTCAATAGC 1588
Db 58331 CACCTTAATAAATTCATTTGGCGATGACAGTCCGATGCAAGCCGCTTCAATAGC 58272
QY 1589 CTATTTGGGCTTTCAAGGGGTCACTTTCAGAGCTGGCGGATCAATATATGSAACATTGA 1648
Db 58271 CTATTTGGGCTTTCAAGGGGTCACTTTCAGAGCTGGCGGATCAATATATGSAACATTGA 58212
QY 1649 CATAGCTGCCAGCGCCATTAGAAAGGAAAGAGGCGCGGAGAGAGAACTGG 1708
Db 58211 CATAGCTGCCAGCGCCATTAGAAAGGAAAGAGGCGCGGAGAGAGAACTGG 58152
QY 1709 TACTGGGACTGGGACTGCACTGAAACGACAGCAGTATCCCGCATTTGGCCGAAAC 1768
Db 58151 TACTGGGACTGGGACTGCACTGAAACGACAGCAGTATCCCGCATTTGGCCGAAAC 58092
QY 1769 TGSAACTGTCAAAAGCCACAGCTAGTATGATTCATTTCAGGGGAGACACTGCGCG 1828
Db 58091 TGSAACTGTCAAAAGCCACAGCTAGTATGATTCATTTCAGGGGAGACACTGCGCG 58032
QY 1829 TGCTCAACGAGCCAAAGTTAGAGGCAATATACATATTTCCAGGCAAGAAACCTTT 1888
Db 58031 TGCTCAACGAGCCAAAGTTAGAGGCAATATACATATTTCCAGGCAAGAAACCTTT 57972
QY 1889 AGGACCTAATATTCACACAGGATGATGACAGAGGACAGGACAGCAGTGTATCACTGCC 1948
Db 57971 AGGACCTAATATTCACACAGGATGATGACAGAGGACAGGACAGCAGTGTATCACTGCC 57912
QY 1949 CAGACCTGCGCTGAATTCATTAAGCTGGGACACAGCTCAACGCGCGGATCGGTTAC 2008
Db 57911 CAGACCTGCGCTGAATTCATTAAGCTGGGACACAGCTCAACGCGCGGATCGGTTAC 57852
QY 2009 ACCACGCGCGCGGAGAGGAGATCTCACTCAAGCAGAGATCGAGATCCACTC 2068
Db 57851 ACCACGCGCGCGGAGAGGAGATCTCACTCAAGCAGAGATCGAGATCCACTC 57792
QY 2069 GCCAAGCAAAACTGAGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2128
Db 57791 GCCAAGCAAAACTGAGGGATGATGATGATGATGATGATGATGATGATGATGATGATG 57732
QY 2129 TGGGCGCATACCCCAAAAGGCTCGGAAATACCCCGAAGCAGGACAGCTTAAAGGAAAC 2188
Db 57731 TGGGCGCATACCCCAAAAGGCTCGGAAATACCCCGAAGCAGGACAGCTTAAAGGAAAC 57672
QY 2189 GAGTCAACTAGATCAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 2248
Db 57671 GAGTCAACTAGATCAAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 57612
QY 2249 GTCTCGACGTACGCCCGAGAGAAACACCCGATTAAGACATTCGCGCGCGGACGATGCT 2308
Db 57611 GTCTCGACGTACGCCCGAGAGAAACACCCGATTAAGACATTCGCGCGCGGACGATGCT 57552
QY 2309 GCGGCAATCCCGTGAAGAGCAATACATTAAGGCGTGGAGATGATGATGATGATGATGATG 2368
Db 57551 GCGGCAATCCCGTGAAGAGCAATACATTAAGGCGTGGAGATGATGATGATGATGATGATG 57492
QY 2369 TGTTCGCGGAGCTACATGCGAGCGCGGAGATCGGATGAGGAAAGCTCATGAGGAG 2428
Db 57491 TGTTCGCGGAGCTACATGCGAGCGCGGAGATCGGATGAGGAAAGCTCATGAGGAG 57432
QY 2429 TGATGAGTGTATGATCACTCCAGGCGCAAGTTCTGTGTTCAGAAACGATCTTTGA 2488

Db 57311 TGAATAGTCTAATGATCATGACTCGAGGCCAGTTCTGTGTGTAGAACATCTTTTGA 57312
Qy 2489 CTCAGACTACACTAGAGGTAAATTAATTCGAATCACTACTAGCGGAGCCACACCCGCTT 2548
Db 57312 CTCAGACTACACTAGAGGTAAATTAATTCGAATCACTACTAGCGGAGCCACACCCGCTT 57312
Qy 2549 GGACTGGAGCAGCAGCGCGCGCCATTGTGACACATCGACAGATTAATTCAGTTCTGCGC 2608
Db 57311 GGACTGGAGCAGCAGCGCGCGCCATTGTGACACATCGACAGATTAATTCAGTTCTGCGC 57312
Qy 2609 ATGACCGATTGTCGTGAAAGGAAGATGAGCTGTGACGCTTCAACACAGTATTCGCGCGA 2668
Db 57351 ATGACCGATTGTCGTGAAAGGAAGATGAGCTGTGACGCTTCAACACAGTATTCGCGCGA 57352
Qy 2669 TGAAGAGAGCTCACTCAGCAGCACTTAATGCTTGTGACATGTTCCGCAAGATGTT 2728
Db 57391 TGAAGAGAGCTCACTCAGCAGCACTTAATGCTTGTGACATGTTCCGCAAGATGTT 57392
Qy 2729 CATGATACACACACCAAGGTGTACTCATCTGTTTGTGACACGCTCACCGAATGATTTCT 2788
Db 57331 CATGATACACACACCAAGGTGTACTCATCTGTTTGTGACACGCTCACCGAATGATTTCT 57332
Qy 2789 GGTTCATACCGAACCAAACTTACCAAAATGCTCTTCAATCTTGTGAGCGGCTGTTCAA 2848
Db 57071 GGTTCATACCGAACCAAACTTACCAAAATGCTCTTCAATCTTGTGAGCGGCTGTTCAA 57072
Qy 2849 CAATTTGGGACCGCATCTACTCAATTTGATGACACAGCAATTTGGAAGACGCTACAGGT 2908
Db 57014 CAATTTGGGACCGCATCTACTCAATTTGATGACACAGCAATTTGGAAGACGCTACAGGT 57015
Qy 2909 GGTTCACGAATATTTTCCCGACGAGTTGACAGCTTAAAGACTATTTAGAAATCATATCGCA 2968
Db 56954 GGTTCACGAATATTTTCCCGACGAGTTGACAGCTTAAAGACTATTTAGAAATCATATCGCA 56955
Qy 2969 TTCTACTCAAAAGCCCACTACCAAGACGCGCATGCGCATTTCTGCGCTCTTACGAGATCT 3028
Db 56894 TTCTACTCAAAAGCCCACTACCAAGACGCGCATGCGCATTTCTGCGCTCTTACGAGATCT 56895
Qy 3029 GGGCAATACACTACTGTAAAGAGAGGACTTTCCAGGACACAGAGCCAGGCGCTGAGGCG 3088
Db 56834 GGGCAATACACTACTGTAAAGAGAGGACTTTCCAGGACACAGAGCCAGGCGCTGAGGCG 56835
Qy 3089 GAGCGTCTTAAAGCTGGCCAGCTGGCGCGGATCAGAAATCGATGAGAGCTGCCCTCCA 3148
Db 56774 GAGCGTCTTAAAGCTGGCCAGCTGGCGCGGATCAGAAATCGATGAGAGCTGCCCTCCA 56775
Qy 3149 GGGCAGAGAGCTGCTAGTGGCTGTATATCTGTAATCTCCCGCAATGACCTTTTAACT 3208
Db 56714 GGGCAGAGAGCTGCTAGTGGCTGTATATCTGTAATCTCCCGCAATGACCTTTTAACT 56715
Qy 3209 GGGCAGAGAGCTGCTAGTGGCTGTATATCTGTAATCTCCCGCAATGACCTTTTAACT 3268
Db 56654 GGGCAGAGAGCTGCTAGTGGCTGTATATCTGTAATCTCCCGCAATGACCTTTTAACT 56655
Qy 3269 GCGGCAAAAGCCAAAGTTGCAATTCGAGTCCCAATTCGCTAGTACTCTCCATTGAGAG 3328
Db 56594 GCGGCAAAAGCCAAAGTTGCAATTCGAGTCCCAATTCGCTAGTACTCTCCATTGAGAG 56595
Qy 3329 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGAGGTCATTTGCTGCGCTTCAAGAGCA 3388
Db 56534 TAGCAGTCCCAAGCTTTGCAAGTCCCTGTGAGGTCATTTGCTGCGCTTCAAGAGCA 56535
Qy 3389 CCACCAACCACTTGTGATACAGCTCTAATGTCAGAGTCCCGGAGAGTTCGAGAGCA 3448
Db 56474 CCACCAACCACTTGTGATACAGCTCTAATGTCAGAGTCCCGGAGAGTTCGAGAGCA 56475
Qy 3449 GGAAGTCTCTTTTCTCGAGTGTGACATTCAGACACAACTTGAAGAAGCTGAGAGA 3508
Db 56414 GGAAGTCTCTTTTCTCGAGTGTGACATTCAGACACAACTTGAAGAAGCTGAGAGA 56415
Qy 3509 GATTCGGCACTGCTTGGCGGTCAGTACAGACGCGCTGGCGCCCAATGGCTTCAATGG 3568

Db 56354 GATTCGGCACTGCTTGGCGGTCAGTACAGACGCGCTGGCGCCCAATGGCTTCAATGG 56355
Qy 3569 ACACCTGACATATCAGCATCAGGCGCAACAGATTTGTGTGATCCTCTGTCTTCCAACTTC 3628
Db 56294 ACACCTGACATATCAGCATCAGGCGCAACAGATTTGTGTGATCCTCTGTCTTCCAACTTC 56295
Qy 3629 CAAGCGCAATCGTCGGCCCAACATCAGTCCCAATTCAGTGTGATGACACAAATGAG 3688
Db 56234 CAAGCGCAATCGTCGGCCCAACATCAGTCCCAATTCAGTGTGATGACACAAATGAG 56235
Qy 3689 GCTGATTAATCTGGAGGCGGAAAGAGCACTCAGAAAGCCCAAGTCCCAACTGAGATGC 3748
Db 56174 GCTGATTAATCTGGAGGCGGAAAGAGCACTCAGAAAGCCCAAGTCCCAACTGAGATGC 56175
Qy 3749 CAAGTATCAACGCTCTGATTAATTAATGCTGAAATGAGAGGCTGATCTGGCCAGCA 3808
Db 56114 CAAGTATCAACGCTCTGATTAATTAATGCTGAAATGAGAGGCTGATCTGGCCAGCA 56115
Qy 3809 CTTGATGAGAGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3868
Db 56054 CTTGATGAGAGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 56055
Qy 3869 GTTGCTTCAAGCTCACTTAATACTGAGGATTTGATCAAGGCGGAAATCTGAGCT 3928
Db 55994 GTTGCTTCAAGCTCACTTAATACTGAGGATTTGATCAAGGCGGAAATCTGAGCT 55995
Qy 3929 GGGCAATACACACTTTAGATGATCATGACGATGCTCTTAACTTCTGAGGCGGAGCA 3988
Db 55934 GGGCAATACACACTTTAGATGATCATGACGATGCTCTTAACTTCTGAGGCGGAGCA 55935
Qy 3989 TACGAGCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4048
Db 55874 TACGAGCTGTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 55875
Qy 4049 GCGTCAACATGATGATGATCTTCTGAGGCTGATTTGCTGAGATTCATCTGAGTATCA 4108
Db 55814 GCGTCAACATGATGATGATCTTCTGAGGCTGATTTGCTGAGATTCATCTGAGTATCA 55815
Qy 4109 ACACAGCAAGAGGCTTTGCGGATATGACTCGATGATACCAAGATAGCACCATCTT 4168
Db 55754 ACACAGCAAGAGGCTTTGCGGATATGACTCGATGATACCAAGATAGCACCATCTT 55755
Qy 4169 GCGTCTGATCTGCTCATCAATTTGCAATTCGATTAATTCGAGGCTGATTTCCAC 4228
Db 55694 GCGTCTGATCTGCTCATCAATTTGCAATTCGATTAATTCGAGGCTGATTTCCAC 55695
Qy 4229 AAATTTGTGGCATGCAAGATCTGCTGCAAGTGAACGAAACCATGTTTGGAGATCAC 4288
Db 55634 AAATTTGTGGCATGCAAGATCTGCTGCAAGTGAACGAAACCATGTTTGGAGATCAC 55635
Qy 4289 AGAGGCCCATTTGGAATTTGTTTCCCAATTCGCGGATCGGAGACGACATC 4348
Db 55574 AGAGGCCCATTTGGAATTTGTTTCCCAATTCGCGGATCGGAGACGACATC 55575
Qy 4349 GATGTTTGGCAAAAGCTGGGCTTCTGATCTGCAAGCTGATCTTGTCTGGCGAGGA 4408
Db 55514 GATGTTTGGCAAAAGCTGGGCTTCTGATCTGCAAGCTGATCTTGTCTGGCGAGGA 55515
Qy 4409 GAAAGTCAACCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4468
Db 55454 GAAAGTCAACCGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 55455
Qy 4469 CATGAGAAAGCAGGATCTGATCTGATGAGGAGGAGAAAGCTTCAAAAGAACTCTCGC 4528
Db 55394 CATGAGAAAGCAGGATCTGATCTGATGAGGAGGAGAAAGCTTCAAAAGAACTCTCGC 55395
Qy 4529 GGCATGCTGTCTGATGATGCTGAGAGGCTTAAATGAGATTTCTGCTGCTGCTGCT 4586
Db 55334 GGCATGCTGTCTGATGATGCTGAGAGGCTTAAATGAGATTTCTGCTGCTGCTGCT 55335
Qy 4587 ACAGAGCAGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4646
Db 55274 ACAGAGCAGAGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 55275

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.bgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES
source
1. 182960
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/clone="RP98-2C22"
39171..48280
misc_feature
/notes="polymorphic rno element which does not exist in celera data and in DMR"
/function="polymorphic site"

ORIGIN

Query Match 98.2%; Score 5051.8; DB 2; Length 182960;

Best Local Similarity 99.6%; Pred. No. 0; Mismatches 17; Indels 5; Gaps 3;

Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

29 GTTCCGCGAGAACGAGACGAAATAGACATTGCAACGATGCGCTATCGGAAGCCAGCGA 88
Db 176386 GTTCCGCGAGAACGAGACGAAATAGACATTGCAACGATGCGCTATCGGAAGCCAGCGA 176327
Qy 89 CTTGATGAGCTTTATCCAGCAATATGCCAAGCGGACATGCGTGTGAAGTACAGCTCGC 148
Db 176326 CTTGATGAGCTTTATCCAGCAATATGCCAAGCGGACATGCGTGTGAAGTACAGCTCGC 176267
Qy 149 CGAAGATCTGGTGAATCTCTTGAAGAGACAAATCAATGTCCTGACGAGCATGCGG 208
Db 176266 CGAAGATCTGGTGAATCTCTTGAAGAGACAAATCAATGTCCTGACGAGCATGCGG 176207
Qy 209 ATTCTTATTCAGCGTTTGTATGATGCGATGCGGACGCGACCTTTAAGATTGACAAA 268
Db 176206 ATTCTTATTCAGCGTTTGTATGATGCGATGCGGACGCGACCTTTAAGATTGACAAA 176147
Qy 269 GTTCCCTGAGAGCGTTCTTGGAGCTTAATAAGCGATTGGGCGGATTTAATGCATAC 328
Db 176146 GTTCCCTGAGAGCGTTCTTGGAGCTTAATAAGCGATTGGGCGGATTTAATGCATAC 176087
Qy 329 GGCTACGTTCTGCAATGATGATCGATCGGCTGGAGAGACAGGAGCAAGTCCGCGA 388
Db 176086 GGCTACGTTCTGCAATGATGATCGATCGGCTGGAGAGACAGGAGCAAGTCCGCGA 176027
Qy 389 GAAGCGCAACTTCTGCTGCGGACCTTCATGAGACACAGATGCTTCCGCCAGCGCT 448
Db 176026 GAAGCGCAACTTCTGCTGCGGACCTTCATGAGACACAGATGCTTCCGCCAGCGCT 175967
Qy 449 GATGACAAAGTGGCCACTAGCTGCTTACGACACAGAACCCAAAGTGGCGGAGAGTT 508

Db 175966 GATGACAAAGTGGCCACTAGCTGCTTACGACACAGAACCCAAAGTGGCGGAGAGTT 175907
Qy 509 CTTTACAGCAATTTGAGACGCTTTTCAATGATACGAGACCCAGAGCTTATGTTCCGCT 568
Db 175906 CTTTACAGCAATTTGAGACGCTTTTCAATGATACGAGACCCAGAGCTTATGTTCCGCT 175847
Qy 569 CTATATACCAACAGTTTGTGACCTTTCGAGATTCACCACTTAAATGTAGAGAGCGGC 628
Db 175846 CTATATACCAACAGTTTGTGACCTTTCGAGATTCACCACTTAAATGTAGAGAGCGGC 175787
Qy 629 CATCCAAACGCTAGTGAATTTCTACAGCAATGTAGGGGATGATGGCCCGACCTCCG 688
Db 175786 CATCCAAACGCTAGTGAATTTCTACAGCAATGTAGGGGATGATGGCCCGACCTCCG 175727
Qy 689 TCGCATGAGAGATTTCTGCTCGTGAATTTGGCTATGTGGAGCAAAAGTTCCAGAGT 748
Db 175726 TCGCATGAGAGATTTCTGCTCGTGAATTTGGCTATGTGGAGCAAAAGTTCCAGAGT 175667
Qy 749 CAACAGAGGCTTACTGCTACTCTTCACTTCAAGCCCTTAAACAGAAATGCAATGAGTGG 808
Db 175666 CAACAGAGGCTTACTGCTACTCTTCACTTCAAGCCCTTAAACAGAAATGCAATGAGTGG 175607
Qy 809 CTTGAGAGAGCGGACCAATATTTGGGTTGAGGAGAGCCACCAAGATGATTAAGCGGC 868
Db 175606 CTTGAGAGAGCGGACCAATATTTGGGTTGAGGAGAGCCACCAAGATGATTAAGCGGC 175547
Qy 869 ACTACATCGGCGCTTTCGTATCATCATGCGCCCAAAACCCAAATGCAAGATGACCGG 928
Db 175546 ACTACATCGGCGCTTTCGTATCATCATGCGCCCAAAACCCAAATGCAAGATGACCGG 175487
Qy 929 TGATGCGGCGCGCTTAACATGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 988
Db 175486 TGATGCGGCGCGCTTAACATGGAATTTTGAATTTTGAATTTTGAATTTTGAATTTTGAATTT 175427
Qy 989 GAACATCTTCCAGCTTAAGACATGAGACATCTTCAAGCAATCTAAGCAATCTAAGCAATCTAAG 1048
Db 175426 GAACATCTTCCAGCTTAAGACATGAGACATCTTCAAGCAATCTAAGCAATCTAAGCAATCTAAG 175367
Qy 1049 TGATTAACAGCAGACTGAGAGAAACGTGTGATGCTCTCAAGAAATCAGAGGCTTCT 1108
Db 175366 TGATTAACAGCAGACTGAGAGAAACGTGTGATGCTCTCAAGAAATCAGAGGCTTCT 175307
Qy 1109 CATTCAGCTATCACTACGCGGACGTTTGTCTGTACAGCTTAAGAAATTTGCTGT 1168
Db 175306 CATTCAGCTATCACTACGCGGACGTTTGTCTGTACAGCTTAAGAAATTTGCTGT 175247
Qy 1169 AAGCTTGTGACATCTCTCAAGAGAACTTACATCAAGTGTATCCGAGAGGCTGAT 1228
Db 175246 AAGCTTGTGACATCTCTCAAGAGAACTTACATCAAGTGTATCCGAGAGGCTGAT 175187
Qy 1229 CACCATGCGCTACATGCTTCAAGAGCGTGAAGAAATTAATCTAGATGCTTCTGAGAGAT 1288
Db 175186 CACCATGCGCTACATGCTTCAAGAGCGTGAAGAAATTAATCTAGATGCTTCTGAGAGAT 175127
Qy 1289 TTTGAGACACTGATTAATTAATTAATCAAGAACGCGGAGGTCATTCGCTTCCAC 1348
Db 175126 TTTGAGACACTGATTAATTAATTAATCAAGAACGCGGAGGTCATTCGCTTCCAC 175067
Qy 1349 AATAGCTCTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1408
Db 175066 AATAGCTCTGAAGTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 175007
Qy 1409 AGACACTGTAATCAATCAAGTGAAGACATTAAGTTCACACTGTGTAGCTGATGAT 1468
Db 175006 AGACACTGTAATCAATCAAGTGAAGACATTAAGTTCACACTGTGTAGCTGATGAT 174947
Qy 1469 GCTGCTCTTGAAGAGTGAAGACGAAAGCGTTGGAAGAAATGCAACCTGTAAGGGA 1528
Db 174946 GCTGCTCTTGAAGAGTGAAGACGAAAGCGTTGGAAGAAATGCAACCTGTAAGGGA 174887
Qy 1529 CACCTTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1588
Db 174886 CACCTTAATAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 174827

1589 CTATTGGGCTTTCAGGCGTCACTTTCAGAGCTGCGGATCAAAATATATGCAACTTTGA 1648
174826 CTAATGGGCTTTCAGGCGTCACTTTCAGAGCTGCGGATCAAAATATATGCAACTTTGA 174767
1649 CATACCTCCCGCGCGCTTTGAAAGGGAACGAAGGCGCGCGAGAGAGAGAGAGAGAGAG 1708
174766 CATACCTCCCGCGCGCTTTGAAAGGGAACGAAGGCGCGCGAGAGAGAGAGAGAGAGAG 174707
1709 TACTGGAGCTGGAGCTGCACTGAAACGAGACGACTGATCCCGCATTTGGCGGAACAC 1768
174706 TACTGGAGCTGGAGCTGCACTGAAACGAGACGACTGATCCCGCATTTGGCGGAACAC 174647
1769 TGGAACTCTGCAAAAGCCGCTAGTATGATTCATTTCAGCGGTGAGCACTGGGC 1828
174646 TGGAACTCTGCAAAAGCCGCTAGTATGATTCATTTCAGCGGTGAGCACTGGGC 174587
1829 TGCTCAAGAGCCAAAGTTAGAGCCCAATATCACTTATTTCCAGGCAAAAGAACTTT 1888
174586 TGCTCAAGAGCCAAAGTTAGAGCCCAATATCACTTATTTCCAGGCAAAAGAACTTT 174527
1889 AGGACTTAATATTCACACAGGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1948
174526 AGGACTTAATATTCACACAGGATGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174467
1949 CAGACTCTGCTGAATTCATATAGCGGTGACACACAGCTACACGCGGAGTCGGTTAC 2008
174466 CAGACTCTGCTGAATTCATATAGCGGTGACACACAGCTACACGCGGAGTCGGTTAC 174407
2009 ACCAGGCGCGCGGAG 2068
174406 ACCAGGCGCGCGGAG 174347
2069 GCCAAGCAAAATGAGGAGATCACTAGCGGTGATTAATTAATTCATTCGCGGAGCAC 2128
174346 GCCAAGCAAAATGAGGAGATCACTAGCGGTGATTAATTAATTCATTCGCGGAGCAC 174287
2129 TGGCGCATACCCCAAAAGGCTCCGGAATACCCCGAAGCACAGCCAGCTCTAGGAAAC 2188
174286 TGGCGCATACCCCAAAAGGCTCCGGAATACCCCGAAGCACAGCCAGCTCTAGGAAAC 174227
2189 GAGTCACTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2248
174226 GAGTCACTAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 174167
2249 GTCCTGAGTACGCGCGAG 2308
174166 GTCCTGAGTACGCGCGAG 174107
2309 GGCAGCAATCCCGTGAAGAGAAATATATAGCGGTGAGATGATGAGACAGCAGACTA 2368
174106 GGCAGCAATCCCGTGAAGAGAAATATATAGCGGTGAGATGATGAGACAGCAGACTA 174047
2369 TGTTCCTCGGAGACTATGCGGAGCGGAGATGCGGATGGGTAGAGAGCTCATGAGAGC 2428
174046 TGTTCCTCGGAGACTATGCGGAGCGGAGATGCGGATGGGTAGAGAGCTCATGAGAGC 173987
2429 TGTTCCTCGGAGACTATGCGGAGCGGAGATGCGGATGGGTAGAGAGCTCATGAGAGC 2488
173986 TGTTCCTCGGAGACTATGCGGAGCGGAGATGCGGATGGGTAGAGAGCTCATGAGAGC 173927
2489 CTCAGCTACACTAGAGGTATATATATGAACTACTCATTTAGCGGAGCCACACCGCTT 2548
173926 CTCAGCTACACTAGAGGTATATATATGAACTACTCATTTAGCGGAGCCACACCGCTT 173867
2549 GGCCTGAG 2608
173866 GGCCTGAG 173807
2609 ATCGAGCATTTGCTGAAAG 2668
173806 ATCGAGCATTTGCTGAAAG 173747

2669 TGGAAAGAGCTCACTGAG 2728
173746 TGGAAAGAGCTCACTGAG 173687
2729 CATGATATACACACAG 2788
173686 CATGATATACACACAG 173627
2789 GGTTCATGCGAAGAACTTACAGAAATGCGCTTCACTCTTTGTTGACGCGGTTGTTCAA 2848
173626 GGTTCATGCGAAGAA--CTTACAGAAATGCGCTTCACTCTTTG--TTGACGCGGTTGTTCAA 173570
2849 CAATTTGGGACCCATCTTACTCAATTTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2908
173569 CAATTTGGGACCCATCTTACTCAATTTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 173510
2909 GGTTCAGAAATATTTCCCAAG 2968
173509 GGTTCAGAAATATTTCCCAAG 173450
2969 TTCTACTCAACGCGCACTACCAAG 3028
173449 TTCTACTCAACGCGCACTACCAAG 173390
3029 GGCCTATACATCTTAAAG 3088
173389 GGCCTATACATCTTAAAG 173330
3089 GACGGTCTTAAAG 3148
173329 GACGGTCTTAAAG 173270
3149 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 3208
173269 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 173210
3209 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 3268
173209 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 173150
3269 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 3328
173149 GGCAGAGAGCTGCTAGTGGCGCTGTATTAATCTGATATACCCCGCAATATGACCTTTTACT 173090
3329 TAGCAGTCCCAAGCTTTGCAAAAGTCCCTGTGAGGTCATTTGCTCGCTTCAAGAGCA 3388
173089 TAGCAGTCCCAAGCTTTGCAAAAGTCCCTGTGAGGTCATTTGCTCGCTTCAAGAGCA 173030
3389 CCACCAACCACTTAGATTCAGCTTACTATGATCCAGCTCCCGGAGATCTTCCGAGAGCA 3448
173029 CCACCAACCACTTAGATTCAGCTTACTATGATCCAGCTCCCGGAGATCTTCCGAGAGCA 172970
3449 GGAAGTCTCTTTTCCCTCGAGAGTGAATTCAGACACAACTTTCAGAAAGAGTGGAGAG 3508
172969 GGAAGTCTCTTTTCCCTCGAGAGTGAATTCAGACACAACTTTCAGAAAGAGTGGAGAG 172910
3509 GATCCGCACTGCTTTCGCGGTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3568
172909 GATCCGCACTGCTTTCGCGGTGATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 172850
3569 ACATTTGAGATTAAGATTCAG 3628
172849 ACATTTGAGATTAAGATTCAG 172790
3629 CAAGAGCAATCGTCGCGCAACACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3688
172789 CAAGAGCAATCGTCGCGCAACACTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 172730
3689 GCTGATATATCTGAG 3748
172729 GCTGATATATCTGAG 172670
3749 CAGGTATACGCTTCTGATTAATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3808

Db	172669	CAGGATGATCAGCGTCTCGATTAAATATAGCTGAAAAAGAGACGTGATACCTGGCCAGCAA	172610
QY	3809	CCTGATGAGAGCGAAAGTGGTGGCTGTGGCCTTTGACCGCTAACAAAGGATCAGCCCGTCGA	3868
Db	172609	CCGTATGGAGAGCGGAAGTGGTGGCTGTGGCCTTTGACCGCTAACAAAGGATCAGCCCGTCGA	172550
QY	3869	GTTCGCTTCAGACGTCACTTACTMAACCTGGGGATTGTGATCAAAAGGCGGAAACTGTAGCT	3928
Db	172549	GTTCGCTTCAGACGTCACTTACTMAACCTGGGGATTGTGATCAAAAGGCGGAAACTGTAGCT	172490
QY	3929	GCCCAATAAGCACTTATAGATCGATCATGTGCGAAATGCTGCTTAATCATTTCTGGAGCGGAGCA	3988
Db	172489	GCCCAATAAGCACTTATAGATCGATCATGTGCGAAATGCTGCTTAATCATTTCTGGAGCGGAGCA	172430
QY	3989	TACGACGTGGTCATCGCTGGACCTGACCGTGCTCAGTAAGATTATAGAGAGCAACAAAT	4048
Db	172429	TACGACGTGGTCATCGCTGGACCTGACCGTGCTCAGTAAGATTATAGAGAGCAACAAAT	172370
QY	4049	GGGTCAACAACGTGATCATCTTTTAAGAGCTGATTTTGTCTGAAACATCATCCAGTGTATCA	4108
Db	172369	GGGTCAACAACGTGATCATCTTTTAAGAGCTGATTTTGTCTGAAACATCATCCAGTGTATCA	172310
QY	4109	ACACAGCAAGAGGGCTTGGGGATATPCAGCTGATGATPCAAAGATAGCAACATCCTT	4168
Db	172309	ACACAGCAAGAGGGCTTGGGGATATPCAGCTGATGATPCAAAGATAGCAACATCCTT	172250
QY	4169	GGCTCTGATCTGTCCCATACAATTCGTCAATCCGGTAATTGCGACGGGTGATTTCCAC	4228
Db	172249	GGCTCTGATCTGTCCCATACAATTCGTCAATCCGGTAATTGCGACGGGTGATTTCCAC	172190
QY	4229	AAATTTGTCCGCATCAAGATCTGCTGGAAATGACCGAAACACCATGTGTTCCGAGATCAC	4288
Db	172189	AAATTTGTCCGCATCAAGATCTGCTGGAAATGACCGAAACACCATGTGTTCCGAGATCAC	172130
QY	4289	AGACCCCACTTGGACATGTGTTCGCCCAATCTGGGGCGAATCGGCCGAGACACCGCAATC	4348
Db	172129	AGACCCCACTTGGACATGTGTTCGCCCAATCTGGGGCGAATCGGCCGAGACACCGCAATC	172070
QY	4349	GATGTTTCGCAAGCTCGCGTCTTCTGCAATCGTCAGACTGTACTTTGTTCTGGGCGAGGA	4408
Db	172069	GATGTTTCGCAAGCTCGCGTCTTCTGCAATCGTCAGACTGTACTTTGTTCTGGGCGAGGA	172010
QY	4409	GAAAGTCAAGCCGAAGCTGTCAAGTGTAAATCCACAGCAAGTTAGGCTCTCAACGTGTA	4468
Db	172009	GAAAGTCAAGCCGAAGCTGTCAAGTGTAAATCCACAGCAAGTTAGGCTCTCAACGTGTA	171950
QY	4469	CATCGAAGACGCGGAACTGCATCACTGGGGGAGGAAGCTCTTCAAAAGAACCTCTCCGC	4528
Db	171949	CATCGAAGACGCGGAACTGCATCACTGGGGGAGGAAGCTCTTCAAAAGAACCTCTCCGC	171890
QY	4529	GGCATCGTGTGATGATTTGGCGGGAAGCCCTTAATAGGATTCCTGCG--TCGAGCACACAA	4586
Db	171889	GGCATCGTGTGATGATTTGGCGGGAAGCCCTTAATAGGATTCCTGCGATCAGACCAACAA	171830
QY	4589	ACAGAGACAGACGCGGTTGCTTCCCTCGGCGCTGAGAGAAAGTGAAGAGCGCGGACAT	4646
Db	171829	ACAGAGACAGACGCGGTTGCTTCCCTCGGCGCTGAGAGAAAGTGAAGAGCGCGGACAT	171770
QY	4647	TAAATTAATATATTATTACTTAATCTTAATATTATTATTAATTAATTAATTAATTAATTAAT	4706
Db	171769	TAAATTAATATATTATTACTTAATCTTAATATTATTATTAATTAATTAATTAATTAATTAAT	171710
QY	4707	ACTTGTGTAACGTCTCGATGCATATTTACGAGAGCAAAATGCGGCCGAAAGAAAGTC	4766
Db	171709	ACTTGTGTAACGTCTCGATGCATATTTACGAGAGCAAAATGCGGCCGAAAGAAAGTC	171650
QY	4767	AAGGCGCATCGCCCATCTATATGTAGACAGAAAAACTATTATTATACATAAACGGGAACTTAA	4826
Db	171649	AAGGCGCATCGCCCATCTATATGTAGACAGAAAAACTATTATTATACATAAACGGGAACTTAA	171590
QY	4827	GGAGTAATTCGCAAAAGTGAATTAATTTGATGACCAATCCGCAATCTCACTTTCAC	4886

Dp	171589	CCGAGTAATCGCAAAAGTGAACCTAATAATTGAGCACAOTCCGCAATCCTACTTCAC	171550
Oy	4887	ATCATGTGTAGCTCTTTTACCAGCTCTTAATAATTAAAGCTGATTCGTTTATTGTA	4946
Dd	171529	ATCATGTTAGCTCTTTTACCAGCTCTTAATAATTAAAGCTGATTCGTTTATTGTA	171470
Oy	4947	AGCCATATACGGCTTTATGGAACACGACGCGTCTTTACCTTTTCAGGC	5006
Dd	171469	AGCCATATACGGCTTTATGGAACACGACGCGTCTTTTCACCTTTTCAGGC	171410
Oy	5007	GACTTTGATCCTTATTCGTTTCGCTCACTAAATTCATATTCATTTATTAATACGA	5066
Dd	171409	GACTTTGATCCTTATTCGTTTCGCTCACTAAATTCATATTCATTTATTAATACGA	171350
Oy	5067	TTCGCTCTTAACAACGACTGATTTTGTAATTAATTAATACACAAGCCACAAGA	5126
Dd	171349	TTCGCTCTTAACAACGACTGATTTTGTAATTAATTAATTAATACACAAGCCACAAGA	171290
Oy	5127	GATTAAAATCAATAAAAA	5145
Dd	171289	GATTAAAATCAATAAAAA	171271
RESULT 9			
AEO03593		31357 bp DNA linear INV 11-AUG-2005	
LOCUS			
DEFINITION	Drosophila melanogaster chromosome 3L, section 74 of 83 of the complete sequence.		
ACCESSION	AEO03593 AEO02647 AE014296		
VERSION	AEO03593.2 GI:23094195		
KEYWORDS			
SOURCE			
ORGANISM	Drosophila melanogaster (fruit fly)		
REFERENCE	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 31357)		
AUTHORS	Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.H., Blazer,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabot,R.B., Abbil,I.J.F., Agbayani,A., An,H.J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Bernier,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brockstein,P., Brottier,P., Buttis,K.C., Butts,D.A., Butler,H., Cadiou,E., Center,H., Chandra,T., Cherry,J.M., Chervin,N.L., Hartley,D., Helman,T.J., Davies,P., de Pablos,B., Delcher,A.L., Deng,Z., Maye,A.D., Dew,I., Dietz,S.M., Dodson,K., Doug,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Fertigera,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Gang,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kenshin,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., LaRoche,P., Lei,Y., Levitsky,A.A., Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Matzel,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Murry,D.J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Murphy,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nueskern,D.R., Pacieb,J.M., Palazzolo,M., Plittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Sampton,M., Skupski,M.P., Smith,T., Spier,B., Sprengling,A.C., Stapleton,M., Strong,R., Sun,E., Svrtkas,R., Tecior,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.Y., Wassenaar,D.A., Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodgett, Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.B., Zaveri,U.S., Zhan,M., Zhang,Q., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,		


```

Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
The genome sequence of Drosophila melanogaster
JOURNAL
PUBMED
10731132
2 (bases 1 to 313157)
Celisner, S.E., Wheeler, D.A., Krommiller, B., Carlson, J.W.,
Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E.,
Hodgson, A., George, R.A., Hoskins, R.A., Laverly, T., Muzny, D.M.,
Neelson, C.R., Pacle, J.M., Park, S., Pfeiffer, B.D., Richards, S.,
Sodergren, E.J., Svirskaas, R., Tabor, P.E., Man, K., Stapleton, M.,
Sutton, G.G., Venter, C., Weinstock, G., Scherer, S.E., Myers, E.W.,
Gibbs, R.A. and Rubin, G.M.
Finishing a whole-genome shotgun: release 3 of the Drosophila
melanogaster euchromatic genome sequence
JOURNAL
PUBMED
12537568
3 (bases 1 to 313157)
Mungall, C.J., Matthews, B.B., Campbell, K.S.,
Mirza, S., Crosby, M.A., Kaminker, J.S., Milburn, G.H., Prochuk, S.E.,
Hirdecky, P., Huang, Y., Kaminker, J.S., Milburn, G.H., Prochuk, S.E.,
Smith, C.D., Tupy, J.D., Whitfield, E.J., Bayraktaroglu, L.,
Bertram, B.P., Bertencourt, B.R., Celisner, S.E., de Grey, A.D.,
Drysdale, R.A., Harris, N.L., Richter, J., Russo, S., Schroeder, A.J.,
Shu, S.Q., Stapleton, M., Yamada, C., Ashburner, M., Gelbart, W.M.,
Rubin, G.M. and Lewis, S.E.
Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review
JOURNAL
PUBMED
12537572
4 (bases 1 to 313157)
Kaminker, J.S., Bergman, C.M., Krommiller, B., Carlson, J.,
Svirskaas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E.,
Rubin, G.M., Ashburner, M. and Celisner, S.E.
The transposable elements of the Drosophila melanogaster
euchromatin: a genome perspective
JOURNAL
PUBMED
12537573
5 (bases 1 to 313157)
Celisner, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R.,
Hoskins, R., Stapleton, M., Pacle, J., Park, S., Svirskaas, R.,
Smith, E., Yu, C. and Rubin, G.
Berkeley Drosophila Genome Project
Drosophila melanogaster release 4 sequence
Unpublished
6 (bases 1 to 313157)
Adams, M.D., Celisner, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
7 (bases 1 to 313157)
FlyBase
Direct Submission
Submitted (22-JUL-2005) FlyBase, Harvard University, Biological
Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA
On Sep 17, 2002 this sequence version replaced g1:7296371.
location/Qualifiers
1. 313157
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/chromosome="3L"
/complement(join(3239..3462,3576..34179))
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037036"
/complement(join(3239..3462,3576..34179))
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/product="CG10586-RA"
/db_xref="FLYBASE:FBgn0037036"

CDS
complement(join(3239..3462,3576..4179))
/gene="CG10586"
/locus_tag="Dmel CG10586"
/old_locus_tag="CG10586"
/note="CG10586 gene product from transcript CG10586-RA"
/codon_start=1
/product="CG10586-PA"
/protein_id="AA51660.1"
/db_xref="FLYBASE:FBgn0037036"
/translacion="MKRLFLFLAGLLINNHALLHNETIDLKAKIVLPAPVQREV
IGRRVITNATLGGVIVAAKRYENFICGGTILHILVITAAHCEDEARAKAWSDGCI
SRSEKIRARQVGFVIAKSAQFKNVTMMDAVILNRPVGMKNGITSLTLPQ
TMVSGKNTNPDDEGKHLRTVSPVIEKICREKRSYSISDSMPCASVIGKMD
ACTYDGGPLVYKQVGVISFGICASRRYPGVYDVHVKPVIYAGTALSRSR"
<4431..>7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037037"
<4431..>7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/product="CG10588-RA"
/db_xref="FLYBASE:FBgn0037037"
4431..7607
/gene="CG10588"
/locus_tag="Dmel CG10588"
/old_locus_tag="CG10588"
/note="CG10588 gene product from transcript CG10588-RA"
/codon_start=1
/product="CG10588-PA"
/protein_id="AA51661.2"
/db_xref="FLYBASE:FBgn0037037"
/translacion="MCFRPLGLRRITRKQVQKALNRQYPMSTASVYVQVHEPPKSD
GDKLYALTLNSGLRAMLIDSYSIDPSSIHRSRSLNSTENFNGKALCAVLVGY
GSFSEPOQYOGAHFVEHMIFMGSEKPEVNEPDSPTKSGGPNATHEETCFYE
LDQTHLDRGMDLPMNLAKAPLMLPDAMSRRSAVQSFQETHMDEAVRQDILSLAS
EGYPHGFSKGNKTKLKEGVDSILHKLKFTYDRHVSARMVVALQAOISLDEL
LVNRCAIDIPSSONSIDVSOYNOKARPOFYDVPLVQVDEVCXELTMVLPMPN
FYRSKPMFTISQILGIVGSCALYLRHLMCISVAVGAAESSDSNISYIPNICIY
LSDDGFHIDEVLEATFAWKLINSQLODSTYKESQIENNNRFOIIPPIDHVS
IVESFNILPSKDVLTGPOLFOYEBESAILRQNMKNFNINIMISSYIPKKNYDOR
EPMNGQFKTISMPKQTMBOGSAITKELAHYQPNPFTTDFKIHVAGSKPVRSR
KPLIRNDICELMFRONIRKLPDGYINLFIPIPVESYKQWLVGLFTVLFVRMA
EQLYPALEAGLTGLYIGDKGLVNRVSGVNEKPLIYVEILINMOTIELDIOVNAFK
DLKKQIYNALINGKSLNDLRSLILENKRFSMTSKYESVDITMDIKSKFENPHK
MYVKGILQGNFTEDQATDLQKYLIDTYKSEKLDNLSLDNHLQIPLGSYLLAKFTLN
EDDSNTIITNYOIGSPDLMECIMDLVEIVSEFPFNQRTQDQVRSQILHRIQY
GLAFLITNTQETKRVADYVEORIEAFRSMELVQMSDTEKNIKRETLINGKILG
GVISDEVILNMSSEIYTKVEFNRRIQIOMLSHITKEDYINFLNDNDKNNLRKLSVG
VGNHNTNSTQAQSGSLSDLDVQDMANETDSAKEMTDKILFLGSEDDPS
SIKDIYAFKKDLVFPLENTNPLANKS"
complement(7844..8836)
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/map="78A1-78A1"
/db_xref="FLYBASE:FBgn0037038"
complement(join(7844..8100,8177..8836))
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/product="CG11037-RA"
/db_xref="FLYBASE:FBgn0037038"
complement(join(7880..8100,8177..8834))
/gene="CG11037"
/locus_tag="Dmel CG11037"
/old_locus_tag="CG11037"
/note="CG11037 gene product from transcript CG11037-RA"

```

```

/codon_start=1
/product="CG11037-PA"
/protein_id="AAF51662.2"
/db_xref="GI:23094197"
/db_xref="FLYBASE:FBgn0037038"
/translation="MEMHLVCCVLACTLVSTQVYGGDEKIVLSLPESQPGKNTLT

```

Query Match 98.2%; Score 5051.8; DB 2; Length 313157;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 5097; Conservative 0; Mismatches 17; Indels 5; Gaps 3;

```

QY 29 GTTCCCGCAGAAAGAGAGAGATAGACATTGCAACGATGGCTTATGGAAAGCCACCGA 88
Db 225546 GTTCCCGCAGAAAGAGAGAGATAGACATTGCAACGATGGCTTATGGAAAGCCACCGA 225605
QY 89 CCTGATGAGCTTTATCCAGCAAAATGCGCAAGCGGAGATGCGTGTGAAGTACAGCTCGC 148
Db 225606 CCTGATGAGCTTTATCCAGCAAAATGCGCAAGCGGAGATGCGTGTGAAGTACAGCTCGC 225665
QY 149 CGAGGATCTGATGACATTCTTACGACGACGACAACTCAATTGTCTGACGACATGGG 208
Db 225666 CGAGGATCTGATGACATTCTTACGACGACGACAACTCAATTGTCTGACGACATGGG 225725
QY 209 ATTCTTATGACGCTTTGATGACATGCGTGAAGCGGACGACCTTTAAGATTGACAAAA 268
Db 225726 ATTCTTATGACGCTTTGATGACATGCGTGAAGCGGACGACCTTTAAGATTGACAAAA 225785
QY 269 GTCCCTGAGAGCGCTTCTCGAGCTAATTAAGCGATTGGGCGAGCTTTTATGATACAC 328
Db 225786 GTCCCTGAGAGCGCTTCTCGAGCTAATTAAGCGATTGGGCGAGCTTTTATGATACAC 225845
QY 329 GGCCTACGTTCTGCGACATGATGATCGGCTGGGAGACAGCAGGAGCACAGTCCGCA 388
Db 225846 GGCCTACGTTCTGCGACATGATGATCGGCTGGGAGACAGCAGGAGCACAGTCCGCA 225905
QY 389 GAAGCGCCAACTTCTGCGCGCACCTTCATGAGACACAGAGTCTTCCGCCACAGGCGCT 448
Db 225906 GAAGCGCCAACTTCTGCGCGCACCTTCATGAGACACAGAGTCTTCCGCCACAGGCGCT 225965
QY 449 GATGCAACAGCTGGCGCACTAGCTTCAAGCACAAAGCGCCAGGTGGCGGAGGATT 508
Db 225966 GATGCAACAGCTGGCGCACTAGCTTCAAGCACAAAGCGCCAGGTGGCGGAGGATT 226025
QY 509 CCTTACAGCATTTGTAACGCTCTTCATGATGAGGACCCAGAGCTTATGTTCCGCT 568
Db 226026 CCTTACAGCATTTGTAACGCTCTTCATGATGAGGACCCAGAGCTTATGTTCCGCT 226085
QY 569 CTATATACCAACGATTGTCACCTTCGCGAGATCCACAGTTAATGTGAGGAGCGCGC 628
Db 226086 CTATATACCAACGATTGTCACCTTCGCGAGATCCACAGTTAATGTGAGGAGCGCGC 226145
QY 629 CATCCAAACGCTAGTGAATCTTCAAGCAGTGTAGGGGATCGATTGGGCCAGACCTCCG 688
Db 226146 CATCCAAACGCTAGTGAATCTTCAAGCAGTGTAGGGGATCGATTGGGCCAGACCTCCG 226205
QY 689 TCGCATGAGCAGATTCTGCTCTGCGCTGGAATTGCGTATGTGAGGCAAAAGTTGCAACAGT 748
Db 226206 TCGCATGAGCAGATTCTGCTCTGCGCTGGAATTGCGTATGTGAGGCAAAAGTTGCAACAGT 226265
QY 749 CAAACAGAGAGGTTACTGCTACTTCAGCCCTTAAAAACAGAAATGGAATGAGATGGG 808
Db 226266 CAAACAGAGAGGTTACTGCTACTTCAGCCCTTAAAAACAGAAATGGAATGAGATGGG 226325
QY 809 CTGGAAGAGGCGCAATATTGGGTTGAGGAGAGCGACCCAGCCAGATGATTAAGCGGCGC 868
Db 226326 CTGGAAGAGGCGCAATATTGGGTTGAGGAGAGCGACCCAGCCAGATGATTAAGCGGCGC 226385
QY 869 ACTAGACTCGGCGCTTCTGCTATCATCTGCGCCAAACCCAAATGTGAACGATGTGACCGG 928
Db 226386 ACTAGACTCGGCGCTTCTGCTATCATCTGCGCCAAACCCAAATGTGAACGATGTGACCGG 226445
QY 929 TGATGCGGCGCGCTTAACATGGAATCTTTGGAATCTAGCTTTGAGGTGCTCCCGCAATT 988

```

```

Db 226446 TGATGCGGCGCGCTTAACATGGAATCTTTGGAATCTAGCTTTGAGGTGCTCCCGCAATT 226505
QY 989 GAACATCTTCCACGCTTAAGAGACATGAGACATATCTAACAGCAAGTACTAGTATCATCAG 1048
Db 226506 GAACATCTTCCACGCTTAAGAGACATGAGACATATCTAACAGCAAGTACTAGTATCATCAG 226565
QY 1049 TGATTAACACGACATGAGGAGAAACGTGTGATGCTCTCAAGAAATCATGAGGCAATTGCT 1108
Db 226566 TGATTAACACGACATGAGGAGAAACGTGTGATGCTCTCAAGAAATCATGAGGCAATTGCT 226625
QY 1109 CATTCTACGCTATCACTACGCGCGAGTTGTGCGCTGTACAGCTTAAAGAAATTTGCGTT 1168
Db 226626 CATTCTACGCTATCACTACGCGCGAGTTGTGCGCTGTACAGCTTAAAGAAATTTGCGTT 226685
QY 1169 AAGCTTCGTGAACATCTTCAAGAGAGAACTACGATCAGAGTGATCCGCGAGCGTGCAAT 1228
Db 226686 AAGCTTCGTGAACATCTTCAAGAGAGAACTACGATCAGAGTGATCCGCGAGCGTGCAAT 226745
QY 1229 CACCATGCGCTTACATGTCTTAAGACGCTGAGAAATTAACATGATGCTTCTGTGAGGACAT 1288
Db 226746 CACCATGCGCTTACATGTCTTAAGACGCTGAGAAATTAACATGATGCTTCTGTGAGGACAT 226805
QY 1289 TTTGAGACACCTGATTAATTTAATACAGAAACAGCGGAGATCATTCGCTTCCAC 1348
Db 226806 TTTGAGACACCTGATTAATTTAATACAGAAACAGCGGAGATCATTCGCTTCCAC 226865
QY 1349 AATAGCTCTGAAGTATATCATTTAAGTATAACATGACACCAAGCTGCTCAAGATCTACAC 1408
Db 226866 AATAGCTCTGAAGTATATCATTTAAGTATAACATGACACCAAGCTGCTCAAGATCTACAC 226925
QY 1409 AGACATCTGATCAATCAAAATGCGAAGACATTAAGTCCACATGTGTGAGCTGATGTGT 1468
Db 226926 AGACATCTGATCAATCAAAATGCGAAGACATTAAGTCCACATGTGTGAGCTGATGTGT 226985
QY 1469 GCGTCTCTTGAAGAGTGGCAGACGAAAGCGTTGGAAGAAATGCCAGCTACTAAGAGA 1528
Db 226986 GCGTCTCTTGAAGAGTGGCAGACGAAAGCGTTGGAAGAAATGCCAGCTACTAAGAGA 227045
QY 1529 CACCTTAAAAAATCATTTAGGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 1588
Db 227046 CACCTTAAAAAATCATTTAGGCGATGCGATGCGATGCGATGCGATGCGATGCGATGCGAT 227105
QY 1589 CTATTTGGCTTTGAGCGCTGACCTTTCCAGAGCTGCGCGATCAAAATATATGAAACATTAGA 1648
Db 227106 CTATTTGGCTTTGAGCGCTGACCTTTCCAGAGCTGCGCGATCAAAATATATGAAACATTAGA 227165
QY 1649 CATAGCTGCGCGCGCGGCAATTGAAAGGGAAGAGAGCGCGCGGAGAGAGAGAACTGG 1708
Db 227166 CATAGCTGCGCGCGCGGCAATTGAAAGGGAAGAGAGCGCGCGGAGAGAGAGAACTGG 227225
QY 1709 TACTGGGACTGGGACTGCACTGAAACGAGACGCACTGTATCCCGCATTTGGCGGAACACC 1768
Db 227226 TACTGGGACTGGGACTGCACTGAAACGAGACGCACTGTATCCCGCATTTGGCGGAACACC 227285
QY 1769 TGAACCTCTGCAAAAGCCCAAGCGCTGATGTGAGATCATTTCAAGCGGTGACACTGCGGC 1828
Db 227286 TGAACCTCTGCAAAAGCCCAAGCGCTGATGTGAGATCATTTCAAGCGGTGACACTGCGGC 227345
QY 1829 TGCTCAAGAGCGCAAGTTAGAGCGCAATATCACTATTTCCAGGCAAAAGAAACCTTT 1888
Db 227346 TGCTCAAGAGCGCAAGTTAGAGCGCAATATCACTATTTCCAGGCAAAAGAAACCTTT 227405
QY 1889 AGGACCTTAATTAATTCACACGAGCATGATGACAGAGCAGACAGCCAGTGTACTGTGCC 1948
Db 227406 AGGACCTTAATTAATTCACACGAGCATGATGACAGAGCAGACAGCCAGTGTACTGTGCC 227465
QY 1949 CAGACTCGCTGATTTCCAAATAGCGGTGCGACACAGCTTAACAAGCCGCGGATTCGTTAC 2008
Db 227466 CAGACTCGCTGATTTCCAAATAGCGGTGCGACACAGCTTAACAAGCCGCGGATTCGTTAC 227525
QY 2009 ACCAGCGCGCGCGGAGAGAGATCTAGTCAACAGCCAGAGATCGAGATCCACCTTC 2068
Db 227526 ACCAGCGCGCGCGGAGAGAGATCTAGTCAACAGCCAGAGATCGAGATCCACCTTC 227585

```

QY 2069 GCCAAGCAGCAAACTGAGGGATCACTAGCTGGTATTGGTAATTATCTATTCGTGGGGCCAC 2128
 Db 227586 GCCAAGCAGCAAAACTGAGGGATCACTAGCTGGTATTGGTAATTATCTATTCGTGGGGCCAC 227645
 QY 2129 TGGGGCCATACCCAAAAGGCTCCGGAATACCCCGAAGCAGCCAGCTCTAGGGAAC 2188
 Db 227646 TGGGGCCATACCCAAAAGGCTCCGGAATACCCCGAAGCAGCCAGCTCTAGGGAAC 227705
 QY 2189 GAGTCACACTAGTACGTAGTGGTGGCTTATGAAAACGAGTATGTAATCTCTACAGGTGGCG 2248
 Db 227706 GAGTCACACTAGTACGTAGTGGTGGCTTATGAAAACGAGTATGTAATCTCTACAGGTGGCG 227765
 QY 2249 GTTCGACGTAACGCCCGAGAGAAACAACCACTAAGACCAATCGCGCCGAGCAGACTGT 2308
 Db 227766 GTTCGACGTAACGCCCGAGAGAAACAACCACTAAGACCAATCGCGCCGAGCAGACTGT 227825
 QY 2309 GGGCCAAATCCGTAAGAGAGAAACATACATTAGGCGTGGAGATGATGACAGCCAGACTA 2368
 Db 227826 GGGCCAAATCCGTAAGAGAGAAACATACATTAGGCGTGGAGATGATGACAGCCAGACTA 227885
 QY 2369 TGTTCGGGGACTACATCGCAGCGGGGGAATGCGGATGGGTAGAAAGCTCATGGAGC 2428
 Db 227886 TGTTCGGGGACTACATCGCAGCGGGGGAATGCGGATGGGTAGAAAGCTCATGGAGC 227945
 QY 2429 TGATGAGTCTGATGACATCGACTCGAGGCCAGTTCTGTGTGTTCAGAACGATCTTTGCA 2488
 Db 227946 TGATGAGTCTGATGACATCGACTCGAGGCCAGTTCTGTGTGTTCAGAACGATCTTTGCA 228005
 QY 2489 CTCACGCTACCTAGAGTAATAATGAACTACTCTATAGCCGAGCCACACCCGCTT 2548
 Db 228006 CTCACGCTACCTAGAGTAATAATGAACTACTCTATAGCCGAGCCACACCCGCTT 228065
 QY 2549 GGACTGAGAGCAGCAGCGGGGCAATTTGACGACATGAGAGCATATTAATTCAGTTCTGCGC 2608
 Db 228066 GGACTGAGAGCAGCAGCGGGGCAATTTGACGACATGAGAGCATATTAATTCAGTTCTGCGC 228125
 QY 2609 ATCGACGATTTGCTGTAAGAGAGATGCGCTGATGACGCTCCACAGATCTTGCGCGA 2668
 Db 228126 ATCGACGATTTGCTGTAAGAGAGATGCGCTGATGACGCTCCACAGATCTTGCGCGA 228185
 QY 2669 TGGAAAGAGCTCACTGACGAGCACTTAATGCGTCTGGACATGTTCCGCAAGATGT 2728
 Db 228186 TGGAAAGAGCTCACTGACGAGCACTTAATGCGTCTGGACATGTTCCGCAAGATGT 228245
 QY 2729 CATGATACACACAGAGGTGTACTCACTGTTTGGACACGCTACCCGAACTGATCT 2788
 Db 228246 CATGATACACACAGAGGTGTACTCACTGTTTGGACACGCTACCCGAACTGATCT 228305
 QY 2789 GGTTCATGCGAAGCAAACTTCACGAAATGCGCTCTTCATCTTTTGAACGCGGTTGTCAA 2848
 Db 228306 GGTTCATGCGAAGCAAACTTCACGAAATGCGCTCTTCATCTTTG-ITGACGCGGTTGTCAA 228362
 QY 2849 CAAATTTGGGACCGATCTACTCAATTTGATGACAGCAAGATTTGGAAGCGCTACAGGT 2908
 Db 228363 CAAATTTGGGACCGATCTACTCAATTTGATGACAGCAAGATTTGGAAGCGCTACAGGT 228422
 QY 2909 GGTTCAGGAATATTTCCCAACGAGTTCAGCTTAAGAACTATTTGATGATCATATGGA 2968
 Db 228423 GGTTCAGGAATATTTCCCAACGAGTTCAGCTTAAGAACTATTTGATGATCATATGGA 228482
 QY 2969 TTCTACTCAACCGCAACTACAGAGCGCGATGCGCATTTCTGGCTTCTTACGGAATCT 3028
 Db 228483 TTCTACTCAACCGCAACTACAGAGCGCGATGCGCATTTCTGGCTTCTTACGGAATCT 228542
 QY 3029 GGCCTAATACATCTGTAAAGAGAGCACTTTCCAGAGACCAAGGCGCTGCGAGCG 3088
 Db 228543 GGCCTAATACATCTGTAAAGAGAGCACTTTCCAGAGACCAAGGCGCTGCGAGCG 228602
 QY 3089 GACGCTCTAAAGCTGAGCCAGCTGCGGCGGATCAGAGTCGATGAGCTGCGCTCCA 3148
 Db 228603 GACGCTCTAAAGCTGAGCCAGCTGCGGCGGATCAGAGTCGATGAGCTGCGCTCCA 228662

QY 3149 GGCAGAGCTGCTATGTGGCCCTGTATTAACCTGAATACCCGCAATGACCTTTTACT 3208
 Db 228663 GGCAGAGCTGCTATGTGGCCCTGTATTAACCTGAATACCCGCAATGACCTTTTACT 228722
 QY 3209 GGCAGACTGCGCAAAAGTATATCAGAGCTGTGCGCATCTGTGATCATTCGACATGAG 3268
 Db 228723 GGCAGACTGCGCAAAAGTATATCAGAGCTGTGCGCATCTGTGATCATTCGACATGAG 228782
 QY 3269 GCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGCTAGTAGCTTCATTTAGAGCA 3328
 Db 228783 GCGGCAAGCCAAAGTTGCAATTCGGGTGCCAATTCGCTAGTAGCTTCATTTAGAGCA 228842
 QY 3329 TAGCACTCCCAAGCCTTTTGCAAGTCCCTCTGTGGGTCCATTTGGCTCGCTTACAGCCA 3388
 Db 228843 TAGCACTCCCAAGCCTTTTGCAAGTCCCTCTGTGGGTCCATTTGGCTCGCTTACAGCCA 228902
 QY 3389 CCACGACCAACTTAGATCAGCTCTACTAGTCACAGCTCCCGGCACTTCTCCGTGAGCA 3448
 Db 228903 CCACGACCAACTTAGATCAGCTCTACTAGTCACAGCTCCCGGCACTTCTCCGTGAGCA 228962
 QY 3449 GAGAGCTGCTCTTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAAGCGTCGAGGA 3508
 Db 228963 GAGAGCTGCTCTTTTCTCGAGCTGACATTCAGCAACAATTCAGAGAAGCGTCGAGGA 229022
 QY 3509 GATCCGGCACTGCTTGGCGGTCAAGTACAGAGCGGCTGGGCCCAATGGCTTCAATGG 3568
 Db 229023 GATCCGGCACTGCTTGGCGGTCAAGTACAGAGCGGCTGGGCCCAATGGCTTCAATGG 229082
 QY 3569 ACACTTGAGTATACAGATCAGAGGCCAAGAGATTCGTGTGATCCCTGCTTCCAACTC 3628
 Db 229083 ACACTTGAGTATACAGATCAGAGGCCAAGAGATTCGTGTGATCCCTGCTTCCAACTC 229142
 QY 3629 CAAGAGCATAGTCGGGCCCAACATCACCACTCAATACCTGAGTCAGCAACAATGAG 3688
 Db 229143 CAAGAGCATAGTCGGGCCCAACATCACCACTCAATACCTGAGTCAGCAACAATGAG 229202
 QY 3689 GCTGATTAATCTGAGCGGGAAAGAACCTCAGAACCGCAAGTCAACCACTGACATGC 3748
 Db 229203 GCTGATTAATCTGAGCGGGAAAGAACCTCAGAACCGCAAGTCAACCACTGACATGC 229262
 QY 3749 CAAGGATCAGCGGTCTGATTAATATGAGCTGAAAATGAGAGCTGTAATCTGCGCAGCA 3808
 Db 229263 CAAGGATCAGCGGTCTGATTAATATGAGCTGAAAATGAGAGCTGTAATCTGCGCAGCA 229322
 QY 3809 CCTGATGAGAGCGAAGTGTGCTGTGAGCTTGAACGCTAACAAAGATCAACCCGTCGA 3868
 Db 229323 CCTGATGAGAGCGAAGTGTGCTGTGAGCTTGAACGCTAACAAAGATCAACCCGTCGA 229382
 QY 3869 GTTGCTTCAGAGTCACTTAACCTGGGGATTTTGATCAAGGGCGGAAACTGTGAGCT 3928
 Db 229383 GTTGCTTCAGAGTCACTTAACCTGGGGATTTTGATCAAGGGCGGAAACTGTGAGCT 229442
 QY 3929 GCCAATTAAGCACTTATGATGATCAGTGAATGCGCTTAACATTTCTGAGGCGGAGCA 3988
 Db 229443 GCCAATTAAGCACTTATGATGATCAGTGAATGCGCTTAACATTTCTGAGGCGGAGCA 229502
 QY 3989 TAGGAGCTGTATCGTGTGCGCTGACAGCTGCTCAGTAAGATTAGAGGACAACAAAT 4048
 Db 229503 TAGGAGCTGTATCGTGTGCGCTGACAGCTGCTCAGTAAGATTAGAGGACAACAAAT 229562
 QY 4049 GCGTCACAACTGATGACATTTCTAGAGCTGATTTTGTCTGAAGATCATCACTGCTATCA 4108
 Db 229563 GCGTCACAACTGATGACATTTCTAGAGCTGATTTTGTCTGAAGATCATCACTGCTATCA 229622
 QY 4109 AACAGAGAAAGAGGCTTTGGGGGATATGACATCGATGATACCAAGGATGAGCAACATCTT 4168
 Db 229623 AACAGAGAAAGAGGCTTTGGGGGATATGACATCGATGATACCAAGGATGAGCAACATCTT 229682
 QY 4169 GCGCTGATCTGTCAATCAATTCATTCATTCGATTAATTCGAGCGGTGAGTTTCCAC 4228
 Db 229683 GCGCTGATCTGTCAATCAATTCATTCATTCGATTAATTCGAGCGGTGAGTTTCCAC 229742
 QY 4229 AATTTGTGCGGCATCAAGATCTCTGTGAAAGTACCAACCACTGTTCCGAGATCAC 4288

[illegible][illegible]

assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center Clone name: KCKK

Center Clone name: CH230-21G3

Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 81073 bases at least Q40

Consensus quality: 86083 bases at least Q30

Estimated insert size: 82180; sum-of-contigs estimation

Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently

consists of 30 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1067: contig of 1067 bp in length

1068 1167: gap of unknown length

1168 2607: contig of 1440 bp in length

2608 2707: gap of unknown length

2708 3856: contig of 1149 bp in length

3857 3956: gap of unknown length

3957 5031: contig of 1075 bp in length

5032 5131: gap of unknown length

5132 6334: contig of 1203 bp in length

6335 6434: gap of unknown length

6435 7914: contig of 1480 bp in length

7915 8014: gap of unknown length

8015 9116: contig of 1102 bp in length

9117 9216: gap of unknown length

9217 10789: contig of 1573 bp in length

10790 10889: gap of unknown length

10890 12215: contig of 1326 bp in length

12216 12315: gap of unknown length

12316 13772: contig of 1457 bp in length

13773 13872: gap of unknown length

13873 15712: contig of 1840 bp in length

15713 15812: gap of unknown length

15813 17948: contig of 2136 bp in length

17949 18048: gap of unknown length

18049 19562: contig of 1514 bp in length

19563 19662: gap of unknown length

19663 21844: contig of 2182 bp in length

21845 21944: gap of unknown length

21945 25551: contig of 3607 bp in length

25552 25651: gap of unknown length

25652 28935: contig of 3285 bp in length

28936 29036: gap of unknown length

29037 30993: contig of 1963 bp in length

31000 31099: gap of unknown length

31100 34406: contig of 3307 bp in length

34507 34506: gap of unknown length

37696: contig of 3190 bp in length

37697 37796: gap of unknown length

40293: contig of 2497 bp in length

FEATURES

source

40294 40393: gap of unknown length
 40394 42819: contig of 2426 bp in length
 42820 42919: gap of unknown length
 42920 45911: contig of 2992 bp in length
 45912 46011: gap of unknown length
 46012 51628: contig of 5617 bp in length
 51629 51728: gap of unknown length
 51729 55917: contig of 4189 bp in length
 55918 56017: gap of unknown length
 56018 59907: contig of 3890 bp in length
 59908 60007: gap of unknown length
 60008 64704: contig of 4697 bp in length
 64705 64804: gap of unknown length
 64805 71041: contig of 6237 bp in length
 71042 71141: gap of unknown length
 71142 78430: contig of 7289 bp in length
 78431 78530: gap of unknown length
 78531 86612: contig of 8082 bp in length
 86613 86712: gap of unknown length
 86713 96091: contig of 9379 bp in length.
 Location/Qualifiers
 1..96091
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-21G3"
 1068..1167
 /estimated_length=unknown
 2608..2707
 /estimated_length=unknown
 3857..3956
 /estimated_length=unknown
 5032..5131
 /estimated_length=unknown
 6335..6434
 /estimated_length=unknown
 7915..8014
 /estimated_length=unknown
 9117..9216
 /estimated_length=unknown
 10790..10889
 /estimated_length=unknown
 12216..12315
 /estimated_length=unknown
 13773..13872
 /estimated_length=unknown
 15713..15812

gap

Query Match 37.4%; Score 1922.4; DB 14; Length 96091;
 Best Local Similarity 94.9%; Pred. No. 0;
 Matches 1934; Conservative 0; Mismatches 102; Indels 1; Gaps 1;

QY 29 GTTGGCGCAGAGCAGATGACATTCGACGATGCTTACGAGCCCGACGA 88
 DB 44850 GTTGGCGCAGAGCAGATGACATTCGACGATGCTTACGAGCCCGACGA 44791
 QY 89 CCTGATGCTTATTCACCAATGCCAGCGCATGCTGTGAAGTACGCTCCG 148
 DB 44790 CCTGATGCTTATTCACCAATGCCAGCGCATGCTGTGAAGTACGCTCCG 44731
 QY 149 CGAGATCTGTGACATTCCTTAGGACGACCAAACTCAATTGCTGACGACATGGG 208
 DB 44730 CGAGATCTGTGACATTCCTTAGGACGACCAAACTCAATTGCTGACGACATGGG 44671
 QY 209 ATTCCCTATTGACGGTTGATGCTGACGCGGACCCACTTTAAGATTGACAAAA 268
 DB 44670 ATTCCCTATTGACGGTTGATGCTGACGCGGACCCACTTTAAGATTGACAAAA 44611
 QY 269 GTCCCTGAGAGCGCTTCTGAGACTAATAAGGATTTGGCAGCGATTTTAATGATAC 328
 DB 44610 GTCCCTGAGAGCGCTTCTGAGACTAATAAGGATTTGGCAGCGATTTTAATGATAC 44551
 QY 329 GGCTACCGTTCTGCCACATGTATGATGATGCGCTGGGAGACAGAGGACACAGTCCGCGA 388

[illegible]

TITLE
JOURNAL
COMMENT

Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,
Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacleab, J.M., Park, S.,
Pfeiffer, B., Poon, L., Segueira, A., Sethi, H., Snir, E.,
Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zieran, L.L. and
Rudin, G.M.

Direct Submission
Submitted (18-FEB-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-111, Berkeley, CA 94720, USA
On Aug 2, 1999 this sequence version replaced gi:5629992.

For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu. All contigs in this submission meet
the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 82 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
571 570: contig of 570 bp in length
651 650: gap of unknown length
1422 1421: contig of 771 bp in length
1502 1501: gap of unknown length
2631 2630: contig of 1129 bp in length
2711 2710: gap of unknown length
3462 3461: contig of 751 bp in length
3542 3541: gap of unknown length
4250 4249: contig of 708 bp in length
4330 4329: gap of unknown length
5406 5405: contig of 1076 bp in length
5486 5485: gap of unknown length
6130 6130: contig of 645 bp in length
6210 6210: gap of unknown length
6964 6964: contig of 754 bp in length
7045 7044: gap of unknown length
7861 7860: contig of 816 bp in length
7941 7940: gap of unknown length
8911 8911: contig of 971 bp in length
8992 8991: gap of unknown length
10372 10371: contig of 1380 bp in length
10452 10451: gap of unknown length
11172 11171: contig of 720 bp in length
11252 11251: gap of unknown length
12640 12640: contig of 1389 bp in length
12721 12720: gap of unknown length
13536 13535: contig of 815 bp in length
13616 13615: gap of unknown length
14247 14246: contig of 631 bp in length
14327 14326: gap of unknown length
15258 15258: contig of 932 bp in length
15339 15338: gap of unknown length
16038 16037: contig of 699 bp in length
16118 16117: gap of unknown length
16763 16763: contig of 646 bp in length
16844 16843: gap of unknown length
17620 17620: contig of 777 bp in length
17701 17700: gap of unknown length
18514 18514: contig of 814 bp in length
18594 18594: gap of unknown length
19276 19276: contig of 682 bp in length
19356 19356: gap of unknown length
19357 19357: gap of unknown length
20244 20244: contig of 888 bp in length
20324 20324: gap of unknown length
20891 20891: contig of 567 bp in length
20972 20971: gap of unknown length
21645 21645: contig of 674 bp in length
21725 21725: gap of unknown length
22681 22681: contig of 956 bp in length
22761 22761: gap of unknown length

22762 22762: contig of 728 bp in length
23490 23490: gap of unknown length
23570 23570: contig of 866 bp in length
24435 24435: gap of unknown length
24515 24515: gap of unknown length
25840 25840: contig of 1325 bp in length
25861 25861: gap of unknown length
25921 25921: contig of 1138 bp in length
27058 27058: gap of unknown length
27138 27138: gap of unknown length
28156 28156: contig of 1018 bp in length
28236 28236: gap of unknown length
28237 28237: contig of 948 bp in length
29185 29185: gap of unknown length
29265 29265: contig of 897 bp in length
30162 30162: gap of unknown length
30242 30242: contig of 1048 bp in length
31290 31290: gap of unknown length
31370 31370: contig of 1035 bp in length
32404 32404: gap of unknown length
32484 32484: gap of unknown length
32485 32485: contig of 1474 bp in length
33958 33958: gap of unknown length
34038 34038: gap of unknown length
35367 35367: contig of 1329 bp in length
35447 35447: gap of unknown length
35448 35448: contig of 1757 bp in length
37205 37205: gap of unknown length
37285 37285: contig of 1882 bp in length
39167 39167: gap of unknown length
39247 39247: contig of 1894 bp in length
41141 41141: gap of unknown length
41221 41221: contig of 1818 bp in length
43039 43039: gap of unknown length
43118 43118: gap of unknown length
44696 44696: contig of 1580 bp in length
44778 44778: gap of unknown length
44779 44779: contig of 2760 bp in length
47538 47538: gap of unknown length
47539 47539: contig of 3418 bp in length
47619 47619: gap of unknown length
51036 51036: contig of 617 bp in length
51037 51037: gap of unknown length
51117 51117: contig of 617 bp in length
51734 51734: gap of unknown length
51813 51813: contig of 621 bp in length
51814 51814: gap of unknown length
52435 52435: contig of 686 bp in length
52515 52515: gap of unknown length
53201 53201: contig of 469 bp in length
53281 53281: gap of unknown length
53749 53749: contig of 446 bp in length
53829 53829: gap of unknown length
54275 54275: contig of 624 bp in length
54355 54355: gap of unknown length
54979 54979: contig of 624 bp in length
55059 55059: gap of unknown length
55681 55681: contig of 622 bp in length
55682 55682: gap of unknown length
55761 55761: contig of 624 bp in length
55762 55762: gap of unknown length
56386 56386: contig of 717 bp in length
56466 56466: gap of unknown length
57183 57183: contig of 550 bp in length
57262 57262: gap of unknown length
57812 57812: contig of 666 bp in length
57892 57892: gap of unknown length
58558 58558: contig of 666 bp in length
58638 58638: gap of unknown length
58659 58659: contig of 724 bp in length
59363 59363: gap of unknown length
59443 59443: contig of 667 bp in length
60109 60109: gap of unknown length
60189 60189: contig of 624 bp in length
60813 60813: gap of unknown length
60893 60893: contig of 580 bp in length
61474 61474: gap of unknown length
61553 61553: contig of 637 bp in length
61554 61554: gap of unknown length
62190 62190: contig of 815 bp in length
62270 62270: gap of unknown length
62271 62271: contig of 815 bp in length
63085 63085: gap of unknown length
63165 63165: contig of 624 bp in length
63166 63166: gap of unknown length
63789 63789: gap of unknown length
63790 63790: contig of 696 bp in length
64565 64565: gap of unknown length

64566 64545: gap of unknown length
 * 64546 65257: contig of 612 bp in length
 * 65337 65337: gap of unknown length
 * 65338 65942: contig of 605 bp in length
 * 65943 66022: gap of unknown length
 * 66023 66659: contig of 637 bp in length
 * 66660 66739: gap of unknown length
 * 66740 67583: contig of 844 bp in length
 * 67584 67663: gap of unknown length
 * 67664 68269: contig of 606 bp in length
 * 68270 68349: gap of unknown length
 * 68350 69001: contig of 652 bp in length
 * 69002 69081: gap of unknown length
 * 69082 69984: contig of 883 bp in length
 * 69985 70044: gap of unknown length
 * 70045 70796: contig of 752 bp in length
 * 70797 70876: gap of unknown length
 * 70877 71572: contig of 696 bp in length
 * 71573 71652: gap of unknown length
 * 71653 72306: contig of 654 bp in length
 * 72307 72386: gap of unknown length
 * 72387 73045: contig of 659 bp in length
 * 73046 73125: gap of unknown length
 * 73126 73949: contig of 824 bp in length
 * 73950 74029: gap of unknown length
 * 74030 74595: contig of 566 bp in length

Query Match 35.7%; Score 1834.4; DB 14; Length 79733;
 Best Local Similarity 95.2%; Pred. No. 0;

Matches 1864; Conservative 0; Mismatches 91; Indels 3; Gaps 2;

QY 1449 ACACGTGTGAGTGTGATGCTGCTCTTTCGAGAGTGSCAGACGAAAGCGTTGAAAGG 1508
 Db 41119 ACACGTGTGAGTGTGATGCTGCTCTTTCGAGAGTGSCAGACGAAAGCGTTGAAAGG 41060
 QY 1509 AATGCGACCGTACTAAGGAGGACCTTAAATAATCCATTGGCGATGAGACCTGCGATGCA 1568
 Db 41059 AATGCGACCGTACTAAGGAGGACCTTAAATAATCCATTGGCGATGAGACCTGCGATGCA 41000
 QY 1569 CGCGCGCATTCAGATACGCTATTGGGCTTTTCAGGCGCTCACTTTCAGAGCTGCGGAT 1628
 Db 40999 CGCGCGCATTCAGATACGCTATTGGGCTTTTCAGGCGCTCACTTTCAGAGCTGCGGAT 40940
 QY 1629 CAAATATATGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAACGAGAGGC 1688
 Db 40939 CAAATATATGAACATTAGACATAGCTGCCACGCGCATTTAGAAAGGAACGAGAGGC 40880
 QY 1689 GCGCGGAGAGGAGAACTGCTACTGGGAGCTGGGACCTGACCTAAACGAGACGCACTGTA 1748
 Db 40879 GCGCGGAGAGGAGAACTGCTACTGGGAGCTGGGACCTGACCTAAACGAGACGCACTGTA 40820
 QY 1749 TCCCGCATTTGGCGGAGCACTGTGAAACCTGGCAAAAGCCACGCTGTATAGATCATT 1808
 Db 40819 TCCCGCATTTGGCGGAGCACTGTGAAACCTGGCAAAAGCCACGCTGTATAGATCATT 40760
 QY 1809 TCAGCGGTGAGCACTCGGCTGCTCAACGAGCCAAAGTTAGAGCGCAATATACCTATAT 1868
 Db 40759 TCAGCGGTGAGCACTCGGCTGCTCAACGAGCCAAAGTTAGAGCGCAATATACCTATAT 40700
 QY 1869 TCCGAGCAAAAGAAACCTTTAGAGCTATATTTCCACACGAGCATGATGACAGAGCA 1928
 Db 40699 TCCGAGCAAAAGAAACCTTTAGAGCTATATTTCCACACGAGCATGATGACAGAGCA 40640
 QY 1929 GCAGCGAGTGTGATCACTGCGCAGACCTGCGCTGAATTCCTCAATGCGGTGACACAGCT 1988
 Db 40639 GCAGCGAGTGTGATCACTGCGCAGACCTGCGCTGAATTCCTCAATGCGGTGACACAGCT 40580
 QY 1989 ACAACGCGGAGTGTGATCACTGCGCAGGCGCGGAGAGCGGACAGAGTATCTCAGTCAAG 2048
 Db 40579 ACAACGCGGAGTGTGATCACTGCGCAGGCGCGGAGAGAGTATCTCAGTCAAG 40520
 QY 2049 CCAGAGTTCAGATCCACTCGGCAACGACAAACCTAGAGGATCAGTACGCTATTGGT 2108

Db 40519 CCAGAGTTCAGATCCACTCGGCAACGACAAACCTAGAGGATCAGTACGCTATTGGT 40460
 QY 2109 AATTACTATCGTGGGCACTGCGGCTATACCCAAAAAGGCTCCGGAATACCCCGAAGC 2168
 Db 40459 AATTACTATCGTGGGCACTGCGGCTATACCCAAAAAGGCTCCGGAATACCCCGAAGC 40400
 QY 2169 ACAGCCAGCTCTTAGGAAACGAGTCCAACTAGTCAAGTGTGTGCTTGAATGAACGAGT 2228
 Db 40399 ACAGCCAGCTCTTAGGAAACGAGTCCAACTAGTCAAGTGTGTGCTTGAATGAACGAGT 40340
 QY 2229 ATGTACTCTACAGGAGCGGGGTCTGACGATAGCGCCGAGAGAAACACAGTAAAGCA 2288
 Db 40339 ATGTACTCTACAGGAGCGGGGTCTGACGATAGCGCCGAGAGAAACACAGTAAAGCA 40280
 QY 2289 TCGGCGCGGCAAGCACTGTGCGCAATCCCGTGAAGCAATACATTAGGCGTGGGA 2348
 Db 40279 TCGGCGCGGCAAGCACTGTGCGCAATCCCGTGAAGCAATACATTAGGCGTGGGA 40220
 QY 2349 GATGATGAGACGACGACATATGTTTCCGAGGACTACATGCGCAGCGCGGATGCGGATG 2408
 Db 40219 GATGATGAGACGACGACATATGTTTCCGAGGACTACATGCGCAGCGCGGATGCGGATG 40160
 QY 2409 GGTAGGAAGCTCATGGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2468
 Db 40159 GGTAGGAAGCTCATGGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 40100
 QY 2469 TGTTCAGAAAGATCTTTCGACTCCAGCTACCTAGAGTAAATGAATGAACTACCTG 2528
 Db 40099 TGTTCAGAAAGATCTTTCGACTCCAGCTACCTAGAGTAAATGAATGAACTACCTG 40040
 QY 2529 AGCGGAGGCAACCGCGCTTGAATGAGACGAGCGCGGCGCATTTTGAACATGAG 2588
 Db 40039 AGCGGAGGCAACCGCGCTTGAATGAGACGAGCGCGGCGCATTTTGAACATGAG 39980
 QY 2589 AGCATATTCAGTTCGCGCATGACGCACTTGGTCTGAAGGAAGATGCGCTGATCAGC 2648
 Db 39979 AGCATATTCAGTTCGCGCATGACGCACTTGGTCTGAAGGAAGATGCGCTGATCAGC 39920
 QY 2649 CTCACACGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCTTCTG 2708
 Db 39919 CTCACACGATATCTGGCCGATGGAAGAGCTCACTCAGCAGCACTTAAATGCTTCTG 39860
 QY 2709 GACATGTTCCGAGATGTTCACTGATATACACACACAGGTGATCTCATGTTTTGAC 2768
 Db 39859 GACATGTTCCGAGATGTTCACTGATATACACACACAGGTGATCTCATGTTTTGAC 39800
 QY 2769 AGGTTCACCGAATGATTTCTGTTTCACTGGAAGAACTTCAAGAAATGCGCTTATCT 2828
 Db 39799 AGGTTCACCGAATGATTTCTGTTTCACTGGAAGAACTTCAAGAAATGCGCTTATCT 39742
 QY 2829 TGTTCGAGCGGTTGTTCAAAATTTGGGCAACGATCTTCAATGATGACAGCAAG 2888
 Db 39741 TGTTCGAGCGGTTGTTCAAAATTTGGGCAACGATCTTCAATGATGACAGCAAG 39683
 QY 2889 AATTGGAAGACGCTACAGTGTGTCAGAAATTTTCCCAAGCAGTTGACGTTAAAGAA 2948
 Db 39682 AATTGGAAGACGCTACAGTGTGTCAGAAATTTTCCCAAGCAGTTGACGTTAAAGAA 39623
 QY 2949 CTAATTGAATATATGGAATTTCTACTCAAAAGCCCACTACCAAGAGCGGATGCGCAT 3008
 Db 39622 CTAATTGAATATATGGAATTTCTACTCAAAAGCCCACTACCAAGAGCGGATGCGCAT 39563
 QY 3009 CTGCGCTTCCTTACGAGTCTGCGCAATCATCTCTTAAAGCAGGACGCTTCCAGCGAC 3068
 Db 39562 CTGCGCTTCCTTACGAGTCTGCGCAATCATCTCTTAAAGCAGGACGCTTCCAGCGAC 39503
 QY 3069 CAGAGCCAGGCTGTGAGCGGAGCGGTCTTAAAGCTGCGCGGATGCGGATGCGGATG 3128
 Db 39502 CAGAGCCAGGCTGTGAGCGGAGCGGTCTTAAAGCTGCGCGGATGCGGATGCGGATG 39443
 QY 3129 TCGATGAGGCTGCGGCTCCAGGCGGAGGCTGCGCTGATGAGGCTGCTGATTAACCTGATAC 3188
 Db 39442 TCGATGAGGCTGCGGCTCCAGGCGGAGGCTGCGCTGATGAGGCTGCTGATTAACCTGATAC 39383

QY 3189 CCGGAAATGACCTTTTACTGCGCCGACCTGCGCAAGATATATCAGACTCTGCCGATCC 3248
DB 39382 CCGGAAATGACCTTTTACTGCGCCGACCTGCGCAAGATATATCAGACTCTGCCGATCC 39323
QY 3349 TGCATTCATTGGACATAGGCGGCGCAAGCCAAAGTTGCAATTGGGTGCCAATTGGCTT 3308
DB 39322 TGCATTCATTGGACATAGGCGGCGCAAGCCAAAGTTGCAATTGGGTGCCAATTGGCTT 39263
QY 3309 AGTAGCTCTGATGAGAGAGTAGAGAGTCCCAAGCTTTGCAAGTCCCTCTGTGGGTCCA 3368
DB 39262 AGTAGCTCTGATGAGAGAGTAGAGAGTCCCAAGCTTTGCAAGTCCCTCTGTGGGTCCA 39203
QY 3369 TTTGCTCTGCTTCAGAGCCACCCACCACTTAGCAT 3406
DB 39202 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNTT 39165

RESULT 12
AR495636 767 bp DNA linear PAT 22-SEP-2004
LOCUS Sequence 596 from patent US 6703491.
DEFINITION AR495636
ACCESSION AR495636
VERSION AR495636.1 GI:52431111
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 767)
AUTHORS Homburger,S.A., Ebens,A.J., Jr., Erickson,C.S., Francis-Lang,H.L.,
Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
Drosophila sequences
Patent: US 6703491-A 596 09-MAR-2004;
JOURNAL Exelixis, Inc.; South San Francisco, CA
FEATURES
source location/Qualifiers
1..767
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 14.2%; Score 732.8; DB 6; Length 767;
Best Local Similarity 98.7%; Pred. No. 9.1e-175;
Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;
QY 2639 CCTGATCAGCTCTCACAGATATCTGGCCGATGAAAGAGCTCACTCAGACCAACTTAA 2698
DB 1 CCTGATCAGCTCTCACAGATATCTGGCCGATGAAAGAGCTCACTCAGACCAACTTAA 60
QY 2699 ATGCGTTCTGACATGTTCCGCAAGATTTTCATGATACACACACCAAGGTGTACTACT 2758
DB 61 ATGCGTTCTGACATGTTCCGCAAGATTTTCATGATACACACACCAAGGTGTACTACT 120
QY 2759 GTTTTGGACAGCGTACCGAAGTGAATCTGTTTCATGCGAAGCAAACTTCACGAATGG 2818
DB 121 GTTTTGGACAGCGTACCGAAGTGAATCTGTTTCATGCGAAGCAAACTTCACGAATGG 178
QY 2819 CTCTTCATCTTGTGACGCGGTTGTTCACAAAATTGGGACCGATCTACTCAATTGCAT 2878
DB 179 CTCTTCATCTTGTGACGCGGTTGTTCACAAAATTGGGACCGATCTACTCAATTGCAT 237
QY 2879 GCACAGCAAGATTTGGAAGAGCTTACAGGTGTCCAGCAATATTTCCCGACGCAATTGCA 2938
DB 238 GCACAGCAAGATTTGGAAGAGCTTACAGGTGTCCAGCAATATTTCCCGACGCAATTGCA 297
QY 2939 GCTTAAAGAACTATTGTAATCATATGGAATTTCTAAGCAAGCCAACTACCAAGAGCGG 2998
DB 298 GCTTAAAGAACTATTGTAATCATATGGAATTTCTAAGCAAGCCAACTACCAAGAGCGG 357
QY 2999 CATGCCATTCCTGGCTCTCTTACGAGATCTGGCCAAATCATACTGTAAAGCAGCACTT 3058
DB 358 CATGCCATTCCTGGCTCTCTTACGAGATCTGGCCAAATCATACTGTAAAGCAGCACTT 417
QY 3059 TCCAGCGACCAAGAGCCAGGCTTGAGCGGAGCGTCTTAAAGCTGCCCAAGCTGGCGGC 3118

DB 418 TCCAGCGACCAAGAGCCAGGCTTGAGCGGAGCGTCTTAAAGCTGCCCAAGCTGGCGGC 477
QY 3119 GGATCAGAAATGATGAGAGTGGCTCCAGGCGCAGAGAGTGGCTAGTGGCCCTGTATTA 3178
DB 478 GGATCAGAAATGATGAGAGTGGCTCCAGGCGCAGAGAGTGGCTAGTGGCCCTGTATTA 537
QY 3179 CTTGAATACCCCGCAATGACCCCTTTACTGCGCGACCTTGCCAAAGGTATATCAGACTC 3238
DB 538 CTTGAATACCCCGCAATGACCCCTTTACTGCGCGACCTTGCCAAAGGTATATCAGACTC 597
QY 3239 TGGCCCATCTGATTCATTCGACATAGAGCGGCGCAAGCCAAAGTTGCAATTGGGTGC 3298
DB 598 TGGCCCATCTGATTCATTCGACATAGAGCGGCGCAAGCCAAAGTTGCAATTGGGTGC 657
QY 3299 CAATTCGCTAGTAGCTCTCATTTGAGCAGTAGACAGTCCCAAGCTTTGCAAGTCCCTC 3358
DB 658 CAATTCGCTAGTAGCTCTCATTTGAGCAGTAGACAGTCCCAAGCTTTGCAAGTCCCTC 717
QY 3359 TGTGGTTCATTTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 3408
DB 718 TGTGGTTCATTTGCTGCTTCCAGAGCCACCAACCACTTAGCATCA 767

RESULT 13
AR510918 767 bp DNA linear PAT 22-SEP-2004
LOCUS Sequence 15878 from patent US 6703491.
DEFINITION AR510918
ACCESSION AR510918
VERSION AR510918.1 GI:52446393
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 767)
AUTHORS Homburger,S.A., Ebens,A.J., Jr., Erickson,C.S., Francis-Lang,H.L.,
Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.
Drosophila sequences
Patent: US 6703491-A 15878 09-MAR-2004;
JOURNAL Exelixis, Inc.; South San Francisco, CA
FEATURES
source location/Qualifiers
1..767
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 14.2%; Score 732.8; DB 6; Length 767;
Best Local Similarity 98.7%; Pred. No. 9.1e-175;
Matches 760; Conservative 0; Mismatches 7; Indels 3; Gaps 2;
QY 2639 CCTGATCAGCTCTCACAGATATCTGGCCGATGAAAGAGCTCACTCAGACCAACTTAA 2698
DB 1 CCTGATCAGCTCTCACAGATATCTGGCCGATGAAAGAGCTCACTCAGACCAACTTAA 60
QY 2699 ATGCGTTCTGACATGTTCCGCAAGATTTTCATGATACACACACCAAGGTGTACTACT 2758
DB 61 ATGCGTTCTGACATGTTCCGCAAGATTTTCATGATACACACACCAAGGTGTACTACT 120
QY 2759 GTTTTGGACAGCGTACCGAAGTGAATCTGTTTCATGCGAAGCAAACTTCACGAATGG 2818
DB 121 GTTTTGGACAGCGTACCGAAGTGAATCTGTTTCATGCGAAGCAAACTTCACGAATGG 178
QY 2819 CTCTTCATCTTGTGACGCGGTTGTTCACAAAATTGGGACCGATCTACTCAATTGCAT 2878
DB 179 CTCTTCATCTTGTGACGCGGTTGTTCACAAAATTGGGACCGATCTACTCAATTGCAT 237
QY 2879 GCACAGCAAGATTTGGAAGAGCTTACAGGTGTCCAGCAATATTTCCCGACGCAATTGCA 2938
DB 238 GCACAGCAAGATTTGGAAGAGCTTACAGGTGTCCAGCAATATTTCCCGACGCAATTGCA 297
QY 2939 GCTTAAAGAACTATTGTAATCATATGGAATTTCTAAGCAAGCCAACTACCAAGAGCGG 2998
DB 298 GCTTAAAGAACTATTGTAATCATATGGAATTTCTAAGCAAGCCAACTACCAAGAGCGG 357

QY	2999	CATGCGCAATTCGCGCTTCCTTAAGGATCTGGCCAAATCAATCTGTAAAGAGGACTT	305
Db	358	CATGCGCAATTCGCGCTTCCTTAAGGATCTGGCCAAATCAATCTGTAAAGAGGACTT	417
QY	3059	TCCGACGCACAGAGCCAGGCGCTGGAGCGAGCGGTCTCTAAAGCTGGGCCAGTGGCGC	3118
Db	418	TCCGACGCACAGAGCCAGGCGCTGGAGCGAGCGGTCTCTAAAGCTGGGCCAGTGGCGC	477
QY	3119	GGATCAGAAAGTCGATGAGAGCTGGCGCTCCAGGCGCAGAGCTGGCTTACTGATTA	3178
Db	478	GGATCAGAAAGTCGATGAGAGCTGGCGCTCCAGGCGCAGAGCTGGCTTACTGATTA	537
QY	3179	CCTGAATAACCCCGCAATAGACCCCTTTTACTGGCGACCTGGCAAAAGTATATCAGAGATC	3238
Db	538	CCTGAATAACCCCGCAATAGACCCCTTTTACTGGCGACCTGGCAAAAGTATATCAGAGATC	597
QY	3239	TGCCGATCTGTCATTCATTTGCGCATGAGGCGGCAAAAGTTGCCAATTCGGGTGC	3298
Db	598	TGCCGATCTGTCATTCATTTGCGCATGAGGCGGCAAAAGTTGCCAATTCGGGTGC	657
QY	3299	CAATTGCGCTATGAGCTCTCCATTGAGAGATGAGTCCCAAGCTTTGGCAAAGTCCTC	3358
Db	658	CAATTGCGCTATGAGCTCTCCATTGAGAGATGAGTCCCAAGCTTTGGCAAAGTCCTC	717
QY	3359	TGTGGGTCGCAATTTGCTGCTTCAGAGCCACACCAACCTTATGACATCA	3408
Db	718	TGTGGGTCGCAATTTGCTGCTTCAGAGCCACACCAACCTTATGACATCA	767

RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	ORIGIN
ARS01058	ARS01058	Sequence 6018 from patent US 6703491.	ARS01058	ARS01058.1	GI:52436533	Unknown.	Unknown.	Unclassified.	1 (bases 1 to 1420)	Homburger,S.A., Ebens,A.J. Jr., Erickson,C.S., Francis-Lang,H.L., Margolis,J.S., Reddy,B.P., Ruddy,D.A. and Buchman,A.R.	Drosophila sequences.	Patent: US 6703491-A 6018 09-MAR-2004;	Exelixis, Inc.; South San Francisco, CA
1	1	1	1	1	1	1	1	1	1	1	1	1	1
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											
Matches 676;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1;									
4465	1	TGTACATCGAAGCAGCGGAACTGCATCATGTGGGGAGGAGAGCTTACAAAGAACTCT	12.8%;	Score 661;	DB 6;	Length 1420;							
Best Local Similarity	99.0%;	Pred. No. 1.9e-156;											

Db	241	CGTTACTGTGTAACTGCTGCATGCAATATTCAGCAGATGCAAAATGCGGCCGAAAAGAA	300
Qy	4763	AGTCAGAGGCGCATGCGCCATCTAAATGTGAGCAGAAAACTATTTTATACATTAACGGGAAC	48222
Db	301	AGTCAGAGGCGCATGCGCCATCTAAATGTGAGCAGAAAACTATTTTATACATTAACGGGAAC	360
Qy	4823	TAAAGCAGATTAATCCGCCAAGTGATCTAAATTTGTAGCCAACTCCGCAATCCCTCATTC	48882
Db	361	TAAAGCAGATTAATCCGCCAAGTGATCTAAATTTGTAGCCAACTCCGCAATCCCTCATTC	420
Qy	4883	TCACATCAGTGTGACGTCTTTTACAGCTCCCTAATTAAGCGATCTGTGTTAAT	4942
Db	421	TCACATCAGTGTGACGTCTTTTACAGCTCCCTAATTAAGCGATCTGTGTTAAT	480
Qy	4943	TGTAAAGCTATATACCGCTCTTTATGGAACCTGAGCGTGCTTTTCTACCTTTGTTTC	5002
Db	481	TGTAAAGCTATATACCGCTCTTTATGGAACCTGAGCGTGCTTTTCTACCTTTGTTTC	540
Qy	5003	AGGCGACTTTGTATCTTTATTTGGTTTCCGCTCACTAAATTCATATTACTATTTAAT	5062
Db	541	AGGCGACTTTGTATCTTTATTTGGTTTCCGCTCACTAAATTCATATTACTATTTAAT	600
Qy	5063	ACGATTTCTGCTCTTACAACTGAAGCTATTTTGTATTAATTTAAATACAAAGCCAC	5122
Db	601	ACGATTTCTGCTCTTACAACTGAAGCTATTTTGTATTAATTTAAATACAAAGCCAC	660
Qy	5123	AAAGATTAATAATCAATPAAAAAA	5145
Db	661	AAAGATTAATAATCAATPAAAAAA	683

[illegible]

```
QY 4703 CGTTACTTGTGTAACTGCTCGATGCAATATATTCAGCAGATGCAAAATCGGCGCGAAAGAA 4762
    |||||
Db 241 CGTTACTTGTGTAACTGCTCGATGCAATATATTCAGCAGATGCAAAATCGGCGCGAAAGAA 300
    |||||
QY 4763 AGTCAAGGGGCGCATGCGCCACTATATGTGAGCAGAAAACTATTATATCATTAACGGGAAC 4822
    |||||
Db 301 AGTCAAGGGGCGCATGCGCCACTATATGTGAGCAGAAAACTATTATATCATTAACGGGAAC 360
    |||||
QY 4823 TAAAGCGAGTAAATCCGCAAGTGTACTAAATTTAGCCAACTCCGCAATCCTCACTTC 4882
    |||||
Db 361 TAAAGCGAGTAAATCCGCAAGTGTACTAAATTTAGCCAACTCCGCAATCCTCACTTC 420
    |||||
QY 4883 TCACATCAGTTGTAGCGTCTTTTACAGCTCTTAATTAACGCTGATTCTGTTAAT 4942
    |||||
Db 421 TCACATCAGTTGTAGCGTCTTTTACAGCTCTTAATTAACGCTGATTCTGTTAAT 480
    |||||
QY 4943 TGTAAAGCTATATACCGCTCTTTATGGAACCTCAGAGTGTGCTTTCTACCTTTGTTTC 5002
    |||||
Db 481 TGTAAAGCTATATACCGCTCTTTATGGAACCTCAGAGTGTGCTTTCTACCTTTGTTTC 540
    |||||
QY 5003 AGCGCACTTGTATCCCTTATTTGTTCCGCTCACTAAATCTAATTACTATTAAT 5062
    |||||
Db 541 AGCGCACTTGTATCCCTTATTTGTTCCGCTCACTAAATCTAATTACTATTAAT 600
    |||||
QY 5063 ACGATTCTGCTCTTACAACTGAACATAATTTTGTAAATTAATTAAATACAGAGCCAC 5122
    |||||
Db 601 ACGATTCTGCTCTTACAACTGAACATAATTTTGTAAATTAATTAAATACAGAGCCAC 660
    |||||
QY 5123 AAAGATTAATCAATTAATAA 5145
    |||||
Db 661 AAAGATTAATCAATTAATAA 683
    |||||
```

Search completed: November 29, 2005, 21:13:39
Job time : 16708 secs

Db 815 TCAATGCTCTGGAGCAGACTTTAACAATGAGCAGATTTGTCACATATATGCACT 874
 Qy 592 TTCTGGAGATCCCGCAGTTAATGAGAGAGCGGCATCCAAACGCTAGTGGAAATCT 651
 Db 875 TACTTGGAGATCCAAACCGCAGTTTCAAGATGACGCAATTAACGCTTAGTGGAAATTT 934
 Qy 652 ACAAGCATGAGGGGATGATGCGCCAGACTTCCTCGCAGTGAAGATGTTCTGCT 711
 Db 935 ACAGCATGTAGAGAAAGTGTGAGGGCAGATCTCAGTAAATAAAGATTTGCGACAGT 991
 Qy 712 CGAAATTTGCTATGTTGAGCAAAAAGTTGACAGGTCAAAACAGAGGCTTACTGCTAC 771
 Db 992 CCCGGTGAATGTAATTTTTCAAAATTTGATGAGTCAAAAATCTGGAAAACATGATAC 1051
 Qy 772 CTTCAGCCCTTAAAAACAGATGGCAGATGAGTGGCTTTGAGCAAGGCC-GACAAATTT 830
 Db 1052 AATTCGAAATGATAAAAATTTCAAGATGAAGATTTCTGAGATGTTAAGACCTTCTCT 1111
 Qy 831 GGGTTGAGGGAGCGACCCAGAGATGATTAAGCGGCCATCACTCGGCCGTTTCGTCA 890
 Db 1112 CTGCTAGTTTACATCATCCAAAGGCTCCACCAAGTTCTCGAGAAACGTTGGAATGGAA 1171
 Qy 891 -----TCACTGCGCCCAAAACCAATGTGAAGATGTGACCGGTATG 933
 Db 1172 CCACCCCGCGCTTGTTCATCCACCCTTGGATCCAAATCTTCAGCTGCAAAAAGAGAG 1231
 Qy 934 CCGGCGCGCTTAACATGGAATCTTTGCAATCTAGCTTGAAGTGGTCCCGCAATTTGAACA 993
 Db 1232 CTGCTGCTGTGATGAGAGAGATTTTATAAGCATTTGATGATGATCTGTAGTACAGA 1291
 Qy 994 TCTTTCACGCTAAGACATGAGACATATCTAACAAGCAATGTAAGTATCATCATGATTA 1053
 Db 1292 TTTATTCAGCGCGAGACCTTGAGGAATCCATTAACAAATTAAGGAAATATATCTGATG 1351
 Qy 1054 AAAACGAGACTGGAGAGAAAGTGTGATGCTCTCAAGAAATCAAGGCATTTGCTCATTC 1113
 Db 1352 ACAAGCATGATTTGGAGAGAGAGATTAATGCTTAAAAAAGATTAGATCTTTACTTTT - 1409
 Qy 1114 TCAGCTATCACACTCAGCGCGAGTTTGTGCTGTACAGCTAAAGAAATGTCTTAAAGCT 1173
 Db 1410 -----GGCTGTGCTGTGATGATGATTAATCTTTCAACATTTGCTTTTGGATGAG 1465
 Qy 1174 TCGTGACATCTCAAGAGAGAACTACATCAAGGTATCCGAGCGCTGATCAACA 1233
 Db 1466 CTTTAACTCTGTCTTAAGACCTCGCGCTCAAGGTATGCGGAGGCTGTATCAAGT 1525
 Qy 1234 TCGCTCATATGCTTAAGACGTTGAGAAATAAATAGATGCTTCTGCTGAGCATTTTGG 1293
 Db 1526 TGGGGCATCTGTCTCATCAAGTTTGGGAAATTAAGTTGACATGAGCTGAAAGCCATTAATGC 1585
 Qy 1294 AGCAGCTGATTAATTTAATAAGAAACAGCGGAAAGGTATTCAGCTTCCCAATTAAG 1353
 Db 1586 CAATATCTTTAATTTAATTTCAAAACAGTGCCTAAATTAATGCAATCTGTGTGTAG 1645
 Qy 1354 CTCTGAAGTATATCAATTAAGTATACATGACCAAAAGCTGCTCAAGATTTCAACAGCA 1413
 Db 1466 CTGTAGGTATTAATTTGCGACACACATCTCCAGGTAAATCTGTCTAATCAAGCA 1705
 Qy 1414 CTCTGAATCATCAAAAGTGAAGACATTAAGTCCACTGTGAGAGTGAATGTGTCTGC 1473
 Db 1706 ---ACTGTACTCTAAGCTGTGTGCAAGTTAAGAGCGCTGTTTGAATTTTAAATTTGC 1762
 Qy 1474 TCTTCGAGAGTGGCAGAGAGCGTTGAAAGAAATGCAACCGTATTAAGGAGACCT 1533
 Db 1763 TTTTACAGAGATGGCAGACATTTCAAGAACAGACATATCATGATTTACTGAAACAA 1882
 Qy 1534 TAAAAAAATCAATTTGGCGATGCAAGCTGCGCATCGCCGCAATTTCCAGATACCCCTATT 1593
 Db 1823 TAAAGAAAGGAATCATGATGCTGATTTCCGAAGCAAGATTAAGAACCGGAAATGTTACT 1882
 Qy 1594 GGGCTTTCAGGCGTCACTTTCCAGAGCTGGGGGATCAAAATATAGAACTTACAGATAG 1653
 Db 1883 GGGGTTTCCAGATCTTCCAGAGAGAGACCTTGTACCAACCTTTGAGATCTCT 1942

Qy 1654 CTGCCAGCGCGC 1666
 Db 1943 CTTACCAAGAAAGC 1955
 RESULT 2
 US-10-104-047-517
 ; Sequence 517, Application US/10104047
 ; Publication No. US20030236392A1
 GENERAL INFORMATION:
 ; APPLICANT: HELIX RESEARCH INSTITUTE
 ; FILE OF INVENTION: NO. US20030236392A1e1 full length cDNA
 ; CURRENT APPLICATION NUMBER: US/10/104,047
 ; PRIOR FILING DATE: 2002-03-25
 ; PRIOR APPLICATION NUMBER:
 ; NUMBER OF SEQ ID NOS: 4096
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 517
 ; LENGTH: 2524
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-104-047-517
 Query Match 2.8%; Score 141.8; DB 6; Length 2524;
 Best Local Similarity 46.1%; Pred. No. 1.1e-31;
 Matches 768; Conservative 0; Mismatches 852; Indels 45; Gaps 7;
 Qy 77 GAAGCCAGCGACTTGATGCTTTATCCAGCAATGCCCAGAGCGGACATGCTGTGA 136
 Db 357 GAGAGCCCGCAGATGAGTACTTGTGCGCCAGTGCAGCAGAGAGACCTCGCGCCG 416
 Qy 137 GGTACAGCTCCCGAGATCTGTGACATTCCTTAGG---AGACACAAATCAATGT 193
 Db 417 GCTGAGGTGCGCAGAGAGTCTGTCTTACCTTGGGCCCCGGGCGCATTCGACCT 476
 Qy 194 CTCACGAGCATGGATTCT-----TATTGACGTTTGTATGTCATGCTGACGG 244
 Db 477 GAGAGAGACCTGGCGCGCTAGGCAAGACATGACAGCGCTCACCGGCTGGGTGCTC 536
 Qy 245 CAGCCACTTATGATGACAAAAGTCCCTGAGGCGTTCTCGAGCTAATAAGCATT 304
 Db 537 GACCAACTACCGGGTATCAATTAATGGGATGAAATTTAAGTCCCTTGTGACAGATT 596
 Qy 305 GGGCAGCGATTTAATGACATACAGGCTACCGTCTGCCCATGTGATCGATCGGCTGG 364
 Db 597 ATCAACAGCTTAAATCTTATGTACAGATGTTATTTAGCTTAAATGACAGATGG 656
 Qy 365 AGACAGAGGAGACAGTCCGCGAGAGAGCGCAACTTGTGCTGCGCACTCATGAGCA 424
 Db 657 AGATGCCAAAGAAAGATGTGAGATGAAGTCAAGCTGATTTGAATTAAGATGATCA 716
 Qy 425 CAGAGTCTTCCGCCAGGCGCTGATTCGAAAGCTGGCCCATAGCTCTTCAAGCA 484
 Db 717 AGTAG-----CACCACTATGTACATTTGGGAGCAGTTGGCTTGTGTTAAACAA 770
 Qy 485 GAAGCGCAAGTGGCGGAGAGATTCCTTCAGAGATTTGAAAGCTCTTATGATAGTACG 544
 Db 771 GAATTTTCATCTCGAAGAAAGCGTGTGTCTTTATGAAACCTTAAACATTTTGG 830
 Qy 545 CACCCAGACCTTATGTTGTGCTGTATATATACAGCAATTTGCACTTCTCGAGATCC 604
 Db 831 GGTCAAGCCACTAGTATACAGAAATTAATGATACCAATTTGTATCTGTTTGAAGCTC 890
 Qy 605 CACAGTTATGTAGAGAGCGCCCATCCAAAGCTAGTGAATATCAACAGATGAGG 664
 Db 891 CAACAGTCAGTGAAGATGCTCAATATTGGCTATAGTGAATTAATGACATGTGG 950
 Qy 665 GATTCGATTTGGCGCAGACCTCGTGCATGACGATGTTCTCGCC-----TCGAAATTGG 720
 Db 951 AGAAAAAGTAGATGATTTTATAGAGAGAAATTTCCCTGCTAGATTAAGAAATGAT 1010

```

QY 721 CTAATGTTGAGCAAAAAGTTGACAGCTGCAACAGAGGGCTCTAGCTTACCTTCAGCCC 780
DB 1011 ATTTGCAAAATTTGATGAAAGTGCAGGCGGATGATTTAGTGTCTGCAAGA 1070
QY 781 TTTAAAAACAGATGCGCATGAGTGGCTTGGACAGAGCCGCAATATTGGTTAGGG 840
DB 1071 TAAAGCTTTCAGATGAAAGAAATCACTGATGAAATAGGCCATTCATCAGCTGCATCAGC 1130
QY 841 AGGAGCCCAACAGATGATTTAAGCGGCACTACCTCGGCCGTTTGG----- 887
DB 1131 CTTCAAGGTTCCCTGCACTTAAACATCCGAAATCTCTCCACAGTGCAGAGAGCCTGG 1190
QY 888 -TCATCACTGCGCCCAAAAACCAATGTGAACGATGTGACCGGTGATGCGGCGCGTAAAC 946
DB 1191 TTCAGAGAGTGGCCCTAAGGTTGAGGTGCTTCTAAGAAAGAGGTGCTGAGACAGTTGA 1250
QY 947 CATGGAATCTTTGGAATCTAGCTTTGAGGTGCTCCCGCAATTTGACATCTTCCAGCTTA 1006
DB 1251 TGAAGATGATTTTATTAAGCTTTTACAGATGTCCTTCTATTCAGATTTATTTCTAGTGC 1310
QY 1007 GGACATGACGATATCTACAGAGAGTACATGATCATCAGTATTAATAAAGCAGACTG 1066
DB 1311 ABAATCTGAAGAAACATTAAATTAATTCAGGGAATTTTGTCTGATGATTAACATGACTG 1370
QY 1067 GGAGAAACGTGTGATGCTCTCAAGAAAGATCAGGGCATTTGCTCATTTCTGAGCTATCAG 1126
DB 1371 GGAATCGGCTCCCAATGCACTGAGAAATTTGATCAGTGTCTT-----GCTGAGC 1424
QY 1127 TCAGCCGAGTTTGTCTGTACAGCTTAAAGAAATTTGCTTAAAGCTTGTGTGAGCAATCT 1186
DB 1425 TGCACAGTATGATGCTTTTTCACATTTTACATTTGATGATGAGCACTTAACTTTTC 1484
QY 1187 CAAGGAGAACTACAGTACAGGTGATCCGCGAGGCGTGCATCAGCATGCGCTACAGTTC 1246
DB 1485 AGCTTAAGATCTTAAATCCAGGTGTGAGAAAGCTTGTATTAATCTGTAGCCCACTTTTC 1544
QY 1247 TAAGAGCTGAGAAATTAATTAAGTCTTCTGTGAGCAATTTTGAAGCACTGATTTAA 1306
DB 1545 AACAGTTTGGGAAACAAGTTTATGATGCGCTGAAAGCATTTGATCACTATTTTAA 1604
QY 1307 TTTAATACGAACAGCGCGAAGTCAATGCTCCGCTTCCACAAATAGCTGTGAAGTAT 1366
DB 1605 TCTCGTCCCAATAGTCAAAAGTCAATGCACTTGTGATGTGACGCAATCAGATTTAA 1664
QY 1367 CATTAAGTATACATGACCAAGAGCTCAAGATCTACAGAGCACTGTGATCAATC 1426
DB 1665 CATTGCGCACTCATGTAACCCAGACTTTAATCTTTAATTAACAGCAATTTGCA--CATC 1721
QY 1427 AAAGTGAAGGACATTAAGTCCACTGTGTGAGTGTGCTGCTCTTTCAGAGAGTG 1486
DB 1722 AAAATAGTTCCTCGTGAAGAGAGTTCATTTGAAATTTTAAATTTTGTGTGAAGAGTG 1781
QY 1487 GGAGAGAAAGCGCTTGAAGAGAGATGCCACCTGATTAAGGACACTTAAAAAATCCAT 1546
DB 1782 GCAAGACTCATCTTGAAGAGAGATGACAGCGCTCTGTGAATTAATTAAGAGGAAT 1841
QY 1547 TGGCGATGAGAGTGGAGTGCAGCGCCCATTTCCAAATAGCCTATTGGGCTTTCAGGGG 1606
DB 1842 TCATGATGCTGAGCGTGAAGCCAGAGTGAAGGCAAAAGACATTAATGAGGCTTTGAAA 1901
QY 1607 TCACCTTTCAGAGCTGCGGATCAAAATATATGGAACATTAGACATAGCTGCCAGAGCCG 1666
DB 1902 CCACCTTCTCTGGAAGCTGAACATTTATTAATTTCCCTTGAAGCCATCTTATCAGAAGAG 1961
QY 1667 ATTAGAAGGAAACGAGAGGCGCGGAGAGGAGGAATCTGGTAC 1711
DB 1962 TCTTCAAACTTAATTAACAGTTCTGCGAGTGTAGCATCTTCC 2006

```

RESULT 3
US-10-104-047-113
Sequence 113, Application US/10104047

```

; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; PRIOR FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 113
; LENGTH: 1908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-104-047-113

Query Match      2.3%  Score 116.6; DB 6; Length 1908;
Best Local Similarity 48.7%; Pred. No. 6,1e-24;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGTGATGCCGGCGCCGTAACATGGAATCTTTCAGATCTAGCTTTGAGGTGTCGCCAA 986
DB 353 GGAGGTGCTGAGAGCATGATGAAGATTTTAAAGCTTTTACAGATGTCCTTCT 412
QY 987 TTGAACATCTTCCACGCTAAGACATGACGATATCTACAGCAAGTACTAGTATC 1046
DB 413 ATTCAGATTTATTTATGTCGAGAACTCGAAGAAATTAATTAATCAGGAAATTTTG 472
QY 1047 AGTATTAATAAGCAGACTGCGAGAAAGTGTGATGCTCTCAAGAGTTCAGGGCATTG 1106
DB 473 TCAGATGATTAACATGACTGGATCAGGTGCATGACGTAAGAAATTTGATCAGT 532
QY 1107 CTCAATTCAGCTATCACACTCAGCGAGTTTGTGCTGTACAGCTAAAGAAATTTGTCG 1166
DB 533 CTGTGTGCTGAG-----CTGCAAGATATATGCTTTTCAACATTTAGCATTTGTG 586
QY 1167 TTAAGCTTGTGAGACATCTCTCAAGAGAACTACATCAAGTATCCGCGAGCTGC 1226
DB 587 GATGAGACATTAATCTTTCAGCTAAGATCTTAATCCAGGTGTGAGAAAGCTTGT 646
QY 1227 ATCAACATCGGCTATAGTCTTAAGAGCGTGAAGAAATTAATTAAGTGCCTTGTGAGAG 1286
DB 647 ATTAATGAGCCCACTTTTCAACATTTTGGAAACAAATTGATCAATGAGCCGTGAAGCC 706
QY 1287 ATTTGAGAGCACTGATTAATTTAATACAGAACAGCGGAGAGTCAATTCGCTTCC 1346
DB 707 ATTTGATCACTACCTTTTATCTGCTCCCAATATGCAAAAGTCAATGCACTTCTGA 766
QY 1347 ACAATAGCTTGAAGTATATCAATTAAGTATACATGACCAAGAGCTGCTCAAGATCTAC 1406
DB 767 TGTGAGCAATCAGATTTATCATCTGCGCATACTCATGTATCCAGACTTATCTTAAATA 826
QY 1407 ACAGACATCTGAATCAATCAAAAGTGCAGAGACATTAAGTCCACACTGTGTAGCGTATG 1466
DB 827 ACMAACAATTGCA---CATCAAAATCAAGTCCCGAGAGAGCGTTCAATTAATTTTA 883
QY 1467 GTGCTGCTCTTTCAGAGAGTGGCAGAGAGCGGTGGAAGAGATGCCACCTACTAAG 1526
DB 884 GATTTATTTGTTGCAAGAGTGGCAGAGCTCATTTATGGAAGACATGAGCCCTTTGTT 943
QY 1527 GACACCTTAATAAATTCATTTGCGATGAGACATGCGATGACGCGCCATTTCCAGATAC 1586
DB 944 GAAATCATTAATAAAGGAATTCATGATGCTGACGCTGAGGCGAGAGTGGAGCAAGAAAG 1003
QY 1587 GCTATTTGGGCTTTCAAGCGTACATTTTCAAGCTGCGGAGTCAATTAATTAAGAACATTA 1646
DB 1004 ACATCAATGAGGCTTAAGAAACCATTTCTGCTGAGAGCTGAACATTAATTAATTCCTT 1063
QY 1647 GACATAGCTGCCAGCGCATTTAAGAAAGGACAGAGGCGGCGAGAGAGAACT 1706
DB 1064 GAGCATCTTATCAGAAAGTCTTCAAACTTAATTAAGAGTTCGGGAGGTATGATCTT 1123

```

QY 1707 GGTAAC 1711
DB 1124 CTTCC 1128

RESULT 4

US-10-205-219-114
Sequence 114, Application US/10205219
Publication No. US2003013803A1
GENERAL INFORMATION:
APPLICANT: Warner-Lambert Company
APPLICANT: Lee, Kevin
APPLICANT: Dixon, Altheair
APPLICANT: Brooksbank, Robert
APPLICANT: Pinnock, Robert
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
FILE REFERENCE: WL-A-018200
CURRENT APPLICATION NUMBER: US/10/205,219
CURRENT FILING DATE: 2002-07-24
PRIOR APPLICATION NUMBER: GB 0118354.0
PRIOR FILING DATE: 2001-07-27
NUMBER OF SEQ ID NOS: 197
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 114
LENGTH: 5614
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: KIAA0627
US-10-205-219-114

Query Match 2.3%; Score 116.6; DB 6; Length 5614;
Best Local Similarity 48.7%; Pred. No. 1.3e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGATATCCCGCCCGCCGTAACCATGGAATCTTTGCAATCTAGCTTTGAGTGTGTCCTCCCA 986
DB 269 GGAGGTGCTGGACGATGTAAGATGATTTAATAAGCTTTACAGATGTCCTTCT 328
QY 987 TTGAACATCTTCCACGCTAGACATGACATGATCTTACAGCAAGTACTAGTATCATC 1046
DB 329 ATTCAATTTATTTCTAGTCCAGAACTGAAAGAAATTAATAAATTAACAGGAAATTTTG 388
QY 1047 AGTATATTAACGACAGATGCGGAAACGTGTGATCTCTTCAAGAAAGATCAGGCAATG 1106
DB 389 TCAGATGATTAACATGATGCGGATGACGCGCAATGCACTGAAGAAATTCATGACTG 448
QY 1107 CTGATTTCTGCTATCATCACTCAAGCCGCACTTGTGCTGTACAGCTTAAGGAATTTG 1166
DB 449 CTGTGTGCTGGAG-----CTGCACAGTATGATTTGCTTTTTCACATTTTACGATTTG 502
QY 1167 TTAAAGCTTGTGACATCTCAAGAGAACTACAGATCAGAGGATCCGCGAGGCGTGC 1226
DB 503 GATGAGCACTTAATTAATTTAGCTTAAGATCTTAAGATCCAGATGTTAGAGAACTTGT 562
QY 1227 ATCAGCATGCTTACATGCTTAAGACGCTGAGAAATTAAGTATGATGCTTCTGCTGAGC 1286
DB 563 ATTACTGTAGCCCACTTTTCAACAGTTTGGGAAACAGTTTGTATGATGCGCTGAAGCC 622
QY 1287 ATTTGGAGCACTGATTAATTTTATACAGAACGCGCGAAGGTCATGATCCGCTTCC 1346
DB 623 ATTGTACTACACTTTTAACTCTGCTCCCAATAGTCAAAAGTCAGGCAACTTCTGGA 682
QY 1347 ACATAGCTGTAAGTATATCATTAATATACATGACATGACCAAGCTGCTCAAGATCTAC 1406
DB 683 TGGGACGATCAGATTTATCTTATGCGCATCTATATCCAGACTTAATCTTTATA 742
QY 1407 ACAGACACTGTGATCAATCAATCAAGTGAAGACATTAAGTCCACATGCTGTGAGCTGAG 1466
DB 743 ACAGACATTTGA---CATCAAAATCAGTTCCCGTGAAGACGTTCAATTTGAATTTTA 799
QY 1467 GTGCTGCTTTCAGAGGTGCGACAGAAAGCGCTTGAAGAAAGATGCAACCTACTAAG 1526

DB 800 GATTTATTTTTCAGAGTGCAGACTCATTTGTAAGAAACATGACGCCGCTTGT 859
QY 1527 GACACCTTAATAAAATTCATTTGCGATGAGATGACGCGCCATTTCCAGATTC 1586
DB 860 GAACTATTAATAAAGGAATTCATGATGCTGACGCTGAGGCGAGTGAAGCAAGAAAG 919
QY 1587 GCCTATTTGGGCTTTCAAGGCGTCACTTTCCAGAGCTGCGCGATCAATATATGAACATTA 1646
DB 920 ACATACATGCGTCTTGAAGAACCACTTCTGCTGTAAGCTGAACATTAATATTCCTT 979
QY 1647 GACATAGCTGCCAGCGGCATTTAGAAAGGAACGAGAGGCGCGAGAGAGAACT 1706
DB 980 GAGCCATCTTATCAGAGAGCTTCAAACTTACTTAAGAGTTCTGCGAGTGAATCT 1039
QY 1707 GGTAAC 1711
DB 1040 CTTCC 1044

RESULT 5

US-10-037-270-324
Sequence 324, Application US/10037270
Publication No. US20030104529A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunding
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhimei
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: PC FL_genes Version 1.0
SEQ ID NO 324
LENGTH: 6445
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4051)
US-10-037-270-324

Query Match 2.3%; Score 116.6; DB 5; Length 6445;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

QY 927 GGATATCCCGCCCGCCGTAACCATGGAATCTTGAATCTTGAAGTGTCCCGCAA 986
DB 362 GGAGGTGCTGAGCACTTATGAAAGTATTTAATAAGCTTTTACAGATGTCCTTCT 421
QY 987 TTGAACATCTTCCAGCCTAAGACATGACGATATCTACAGCAAGTACTAGTATC 1046
DB 422 ATTCAATTTATTTCTAGTGAAGACTCGAAGAAACATTTAATAATCAGGAAATTTTG 481


```

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDV3
CURRENT APPLICATION NUMBER: US/10/122,851
PRIORITY FILING DATE: 2002-04-12
PRIORITY APPLICATION NUMBER: 09/620,312
PRIORITY FILING DATE: 2000-07-19
PRIORITY APPLICATION NUMBER: 09/552,317
PRIORITY FILING DATE: 2000-04-25
PRIORITY APPLICATION NUMBER: 09/488,725
PRIORITY FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 324
LENGTH: 6445
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4051)
US-10-122-851-324

```

Query Match 2.3%; Score 116.6; DB 9; Length 6445;

Best Local Similarity 48.7%; Pred. No. 1.5e-23;

Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

927 GGTGATCCGCGCGCGTAAACATGTAATCTTGCATCTTGAAGTGTCCCGCAA 986
362 GGAGTCTCTGGAGAGTGTGAAGATGATTTTAAAGCTTTAAGATGTCCTTCT 421
987 TTGAACATCTTCAACGCTAAGAGATGAGATCTTCAAGCAAGTACTAGATCATC 1046
422 ATTCAGATTATTTAGTTCGAGAACTCGAAGAACTTAATAATCAAGGAAATTTTG 481
1047 AGTGATTAAGAGAGATGAGAAACGCTGTGATCTCTCAAGAGATCAGGCGATTG 1106
482 TCAGATATTAACATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 541
1107 CTCATCTCAGCTATCAGCTACAGCCGAGTTTGTCTGCTGACAGCTAAAGATTGTC 1166
542 CTGCTTCTGAGAG-----CTGCAAGATGATGCTTTTTCACATTAACATTTGTC 595
1167 TTAAGCTTGTGACATCTTCAAGAGAGAACTACAGATCAAGGTGATCCGAGCGTGC 1226
596 GATGAGACACTTAACATTTAGTCAAGATCTTAGATCCAGGTGTTAGAGAGCTTGT 655
1227 ATCAACATCGCTCAATGCTCAAGAGAGAACTAAGTATGCTTCTGCTGAGC 1286
656 ATTAAGTGTAGCCCACTTCAAGAGTTTGGAGAAAGATTGATCAAGGCTGAAGCC 715
1287 ATTTGAGAGACCTGATTAATTTATACAGAAACGCGAAGCTGATTCGCGCTTCC 1346
716 ATTTGATCTACATTTTATCTGCTCCCAATAGTGCAGAAAGCTGAGCACTTTCGA 775
1347 ACAATAGCTCTGAAGTATATCAATTAACATGATCAAGTCAAGTCAAGATCTAC 1406
776 TGTGACACATCAATATTAATCAATGATGATGATGATGATGATGATGATGATGAT 835
1407 ACAGAGCTCTGAATCAATCAAAATGAGAGACATTAAGTCAAGTCTGAGCTGATG 1466
836 ACAAGCAATTTGCA---CATCAAAATCAAGTCTGAGAGAGAGCTTCAATTTGAT 892
1467 GTGTGCTCTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1526
893 GATTATTTGTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 952
1527 GACACCTTAAAAAAATTCATTTGGAGATGACAGATGAGATGACAGCGCCCATTTCCAGATAC 1586

```

```

DB 953 GAAACTATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1012
QY 1587 GCCTATTGGGCTTTTACAGCGGCTACCTTCCAGAGCTGGCGGATCAATATATGAGACATTA 1646
DB 1013 ACATATATGAGTCTTGAAGAACCACTTCTGAGTGAAGCTGAAACATATATATTCCTT 1072
QY 1647 GACATAGCTCCCGAGCGGCGCATTTAGAAAGGAAACGAGAGAGAGAGAGAGAGAGAG 1706
DB 1073 GAGCCATCTTATCAGAAAGAGTCTTCAAACTTACTTAAGAGTTTGGCAGTGTAGATCT 1132
QY 1707 GGTAC 1711
DB 1133 CTTC 1137

```

RESULT 8

US-10-037-270-323

Sequence 323, Application US/10037270

Publication No. US20030104529A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aidong J.

APPLICANT: Yang, Yonghong

APPLICANT: Wang, Jian-Rui

APPLICANT: Zhou, Ping

APPLICANT: Ma, Yundong

APPLICANT: Wang, Dunhui

APPLICANT: Wang, Zhiwei

APPLICANT: Tillinghast, John

APPLICANT: Dmanac, Radoje T.

TITLE OF INVENTION: No. US20030104529A1e1 Nucleic Acids and

FILE REFERENCE: 784CIP2B

CURRENT APPLICATION NUMBER: US/10/037,270

PRIORITY FILING DATE: 2002-01-04

PRIORITY APPLICATION NUMBER: 09/552,317

PRIORITY FILING DATE: 2000-04-25

PRIORITY APPLICATION NUMBER: 09/488,725

PRIORITY FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1104

SOFTWARE: pc_fl_genes Version 1.0

SEQ ID NO 323

LENGTH: 6487

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (164)..(4093)

US-10-037-270-323

Query Match 2.3%; Score 116.6; DB 5; Length 6487;

Best Local Similarity 48.7%; Pred. No. 1.5e-23;

Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

927 GGTGATCCGCGCGCGTAAACATGTAATCTTGCATCTTGAAGTGTCCCGCAA 986
362 GGAGTCTCTGGAGAGTGTGAAGATGATTTTAAAGCTTTAAGATGTCCTTCT 421
987 TTGAACATCTTCAACGCTAAGAGATGAGATCTTCAAGCAAGTACTAGATCATC 1046
422 ATTCAGATTATTTAGTTCGAGAACTCGAAGAACTTAATAATCAAGGAAATTTTG 481
1047 AGTGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1106
482 TCAGATATTAACATGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 541

```

```

QY 1107 CTCATTCTGACGATATCACTACCGCGAGTTTGTGCTGTACAGCTAAGAAATGTCG 1166
Db 542 CTGTTCTCTGGAG-----CTGCACAGATATGCTTTTTCAACTTAACGATTTG 595
QY 1167 TTAAAGCTTGTGACATCTTCAGAGAAAGAACTACGATCAGAGTGATCCGCGAGCGTGC 1226
Db 596 GATGAGACACTTAAACCTTTCAGCTAAGATCTTGAATCCAGGCGGTAAAGAACTTGT 655
QY 1227 ATCAACATCCGCTCATATGTCTTAAGACGCTGAAGAAATAAAGTAGTGCCTTGTGAGC 1286
Db 656 ATTACTGTAGCCCACTTTCACAGCTTTTGGAAACAAAGTTGATCATGCGCTGAAGCC 715
QY 1287 ATTTTGGAGACCTGATTAATTAAATACAGAACCGCGAAGCTCATTTGATCGCTTCC 1346
Db 716 ATTGTACCTACATTTTAAATCTGCTCCCAATGTCGCAAAAGTCATGGCAACTTGGGA 775
QY 1347 ACATAGCTCTGAAGTATATCATTAATATACATACATGACCAAGCTGTCAAGATCTAC 1406
Db 776 TGTGACAGCAATCAATTTATCATCTGCGCATACCTCATATACCTTAACCTTAAATA 835
QY 1407 ACAGACACTCTGAATCAATCAAAAGTCAGAGACATAAGTCCACACTGTGTGAGCTGATG 1466
Db 836 ACAAGCAATTGCA---CATCAAAATCAGTTCCCTGTGAGAGAGCTTATTTGAATTTTA 892
QY 1467 GTGCTGCTCTTGCAGAGTGGCAGACGAGCGCTTGGAAAGAAATGCGACCTTAAGG 1526
Db 893 GATTTATTTTGGCAAGATGGGAGACTCATTTCAATGGAAAGACATTTGACGCGCTTGTG 952
QY 1527 GACACCTTAAATAATTCATTGGCGATGCAAGATGCGATCAGCGCCCATTCAGATAC 1586
Db 953 GAAACATATTAATAAAGGAATTCATGATGCTGACGCTGAGGCCAGAGTGGGAGGAAG 1012
QY 1587 GCGTATGGGCTTTCAGGCGCTCACTTCCAGAGCTGGCGGATCAATATATGAACTTA 1646
Db 1013 ACATACATGGGCTTTAGAAACCACTTCTGTGAGAGCTGAACATTAATATTCCTT 1072
QY 1647 GACATAGCTGCCAGCGCGCATTAAGAAAGGAACGAGAGCGCGAGAGAGGAAT 1706
Db 1073 GAGCCATCTTATCAGAAAGATCTTCAAACTTAAAGAGTTCTGCAATGTAAGATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTCC 1137

```

RESULT 9

```

US-10-117-722-323
; Sequence 323, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117, 722
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 323
; LENGTH: 6487
; TYPE: DNA
; ORGANISM: Homo sapiens

```

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(4093)
US-10-117-722-323

```

```

Query Match 2.3%; Score 116.6; DB 6; Length 6487;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

QY 927- GGTGATGCCGCGCGCTTAACATGAAATCTTTGAAATCTAGCTTTGAGGTGTCGCCGAA 986
Db 362 GGAAGTGTGAGAGCACTTATGAATGATTTTATTAAGCTTTTACAGATGCTTCT 421
QY 987 TTGAACATCTTCCACGCTAAGACATGAGCATATCTACAGCAAGTACTAGTATC 1046
Db 422 ATTCAATTTATTTATGTCGAGAACTCGAAGAAACATTAATTAATTAATCAAGGAAATTTG 481
QY 1047 AGTGATTAATAACGACACTGGGAGAAACCTGTGATGCTCTCAAGAAATCAGGCAATG 1106
Db 482 TGAATGATTAACATGACTGGGATCAGCGTCCAAATGCACTGAAGAAATTCGATCACTG 541
QY 1107 CTCATTCTGAGCTATCACTACCTCAGCGCGAGTTTGTGCTGTACAGCTAAAGAAATTCG 1166
Db 542 CTGTTGCTGGAG-----CTGCACAGTATGATGCTTTTTCACATTTACGATTTG 595
QY 1167 TTAAAGCTTGTGACATCTCTCAAGAGAACTACATCAGAGTATCCGAGGCGTGC 1226
Db 596 GATGAGCACTTAACTTACCTTACGCTAAGATCTTAAGATCCAGGTGTGAGAAAGCTTGT 655
QY 1227 ATCACCATGCGCTTACATGCTTAAAGCGCTGAAGAAATACTAGATGCTTCTGTGAGC 1286
Db 656 ATTACTGTAGCCCACTTTCACAGCTTTGAAACAAAGTTGATCATGCGCTGAAGCC 715
QY 1287 ATTTTGGAGACCTGATTAATTAAATACAGAACAGCGCAAGGCTTATGCTATCCGCTTCC 1346
Db 716 ATTGTACTACACTTTTATATCTGCTCCCAATATGCAAAAGTCAATGCACTTGTGA 775
QY 1347 ACATAGCTCTGAAGTATATCATTAATATACATGACCAAAAGCTGTCAAGATCTAC 1406
Db 776 TGTGACAGCAATCAGATTAATATCATCTGCGCATATCAATGACCACTTAACTTAAATA 835
QY 1407 ACAGACACTCTGAATCAATCAAAAGTCGAAAGACATTAAGTCCACACTGTGTGAGCTGATG 1466
Db 836 ACAAGCAATTGCA---CATCAAAATCAGTTCCCGTGAAGAGAGCTTCAATTTTGA 892
QY 1467 GTGCTGCTCTTGCAGAGTGGCAGACGAAAGCGCTTGGAAAGAAATGCCAGTACTAAG 1526
Db 893 GATTTATTTTGCAGAGATGGGAGACTCATTTATGGAAGACATGACCGCTTGTG 952
QY 1527 GACACCTTAAATAATTCATTGGCGATGCACTGCGATGACGCGCCATTCAGATAC 1586
Db 953 GAAACATTAATAAAGGAATTCATGATGCTGACGCTGAGGCCAGAGTGAAGCAAAAG 1012
QY 1587 GCGTATGGGCTTTCAGGCGTCACTTTCCAGAGCTGGCGGATCAATATATGGAACATTA 1646
Db 1013 ACATACATGGGCTTTAGAAACCACTTCTGTGAGAGCTGAAGCAATTAATATTCCTT 1072
QY 1647 GACATAGCTGCCAGCGCGCATTAAGAAAGGAACGAGAGCGCGGAGAGAGGAAT 1706
Db 1073 GAGCCATCTTATCAGAAAGATCTTCAAACTTAAAGAGTTCTGCAATGTAAGATCT 1132
QY 1707 GGTAC 1711
Db 1133 CTTCC 1137

```

RESULT 10

```

US-10-122-851-323
; Sequence 323, Application US/10122851
; Publication No. US20050239060A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua

```



```

APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Dimanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDV3
CURRENT APPLICATION NUMBER: US/10/122,851
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 323
LENGTH: 6487
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(4093)
US-10-122-851-323

```

```

Query Match 2.3%; Score 116.6; DB 9; Length 6487;
Best Local Similarity 48.7%; Pred. No. 1.5e-23;
Matches 382; Conservative 0; Mismatches 394; Indels 9; Gaps 2;

```

```

QY 927 GGTATGCGCGCGCTTACCATGATCTTTCAGTCTTGGAGTGTCCCGCA 986
DB 362 GGAGGTGCTGAGCAGTGTATGATGATATTATTAAGCTTTTACAGATGCTCTTCT 421
QY 987 TTGAACATCTTCCAGCTAAGACATGAGCATATCTCAAGCACTAGTATCATC 1046
DB 422 ATTGATATTATTTCTAGTCGAGAACTCGAAGAACTTAATTAATAGGAAATTTG 481
QY 1047 AGTATATTAAGCGAGCTGAGAAAGCTGTGATGCTCTCAAGAGATAGGCAATG 1106
DB 482 TCAATATGATTAACATGATGAGATGAGCTGCAATGACATGAAGAAATTCGATCATG 541
QY 1107 CTCATTTCTACGCTTCACTCAGCGCGAGTTTGTGCTGACAGCTTAAGATTTGCG 1166
DB 542 CTGTGTGCTGAG-----CTGCACAGTATGATGCTTTTTCACATTTAGATGTG 595
QY 1167 TTAAGCTCTGTCGATCTCTCAAGAGAGAACTACAGCTACAGTATCGCGAGCGTGC 1226
DB 596 GATGAGAGACTTAACCTTTCAGCTAAGATCTTATGATCCAGGTGTTAGAGAGCTTGT 655
QY 1227 ATCAACATGCGCTACATGCTPAAGCGCTGAAGAAATTAAGTATGCTTGTGTCGAGC 1286
DB 656 ATTACTGTAGGCCCACTTTCAAGTTTGGGAAACAAGTTGATCATGCGCTGAAGCC 715
QY 1287 ATTTTGGAGCAGCTGATTAATTTAATACAGAACAGCGGAGAGCTTATCGCTTCC 1346
DB 716 ATTGTACCTTACACTTTTAAATCTGCTCCCAATAGTGAAGATCAGCAACTTTCGA 775
QY 1347 ACAATAGCTGAGATATATATTATAGATACATGACCAAGAGCTGCTCAAGATCTAC 1406
DB 776 TGTGCAAGATACAGATTTATCATTCGGGATCTCATGATCCCAAGCTTAATCCTTAATA 835
QY 1407 ACAGACACTGTGAATCAATCAAGTGAAGACATPAAGTCCACACTGTGTAGCTGATG 1466
DB 836 ACAAGCAATTTCA--CATCAAAATCAAGTCCCGTGAAGAGACGTTTCAATTTTGA 892
QY 1467 GTGCTGCTCTGAGGAGTGGCAGAGAGGCGTTGGAAGAGATGCCACGCTACTAAG 1526
DB 893 GATTATTTGTCAGAGGTGCAAGCTTCATTCATTGGAAAGACATGACGCGCTTGTT 952
QY 1527 GACACCTTAAATTAATCATTTGGGATGAGATGCGATGCGACCGCATTCAGATAC 1586
DB 953 GAAATCTTAAATTAAGGAGATTCATGATGCTGACCGTGAAGCCAGATGAGCAAGAAAG 1012

```

```

QY 1587 GCTTATGGGCTTTCAGGCGTCACTTTCAGAGCTGGCGGATCAATATATGAACTTA 1646
DB 1013 ACATACATGGGTCTTAATAAACCACTTTCGTTGAGACTCAAACTTATTAATTCCTT 1072
QY 1647 GACATAGCTGCCAGCGCGCATAGAAAGGAGAGAGAGGCGGAGAGAGAACT 1706
DB 1073 GAGCATCTTATACAGAAAGATCTTCAAACTTACTTAAGAGTTTGGCAGATGATCATCT 1132
QY 1707 GGTAAC 1711
DB 1133 CTTCCT 1137

```

```

RESULT 11
US-10-357-930-59229/c
Sequence 59229, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
TITLE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 59229
LENGTH: 608
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 33
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-59229

```

```

Query Match 1.8%; Score 92; DB 8; Length 608;
Best Local Similarity 51.7%; Pred. No. 1.2e-16;
Matches 237; Conservative 0; Mismatches 215; Indels 6; Gaps 1;
QY 232 CATGCTGAGCGGACCACTTAATGATGCAAAAGTCCCTGAGCGCTTCTCGAGC 291
DB 521 CTGGGTGAGACTTACGAATTAAGAAGTGTGCTGTGGGAGATGACATCTGTCCGCC 462
QY 292 TAAATTAAGGATTTGGGAGGATTTTATGATACACGGCTACCGTTCTGCCATGTGA 351
DB 461 TGGTGACCCCGCTGAGATCGGTTCAAGCGCAGATCGGCAAGTGTCCCAAGTCTAA 402
QY 352 TCGATGGCTGGGAGACAGAGGAGACAGTCCGCGAGAGAGGCAACTTCTGCTGGCG 411
DB 401 TAGACAGACTAGAGATGCTTAAAGACTCTGTGAGGAGACAGACTTAATCTGTGCTTA 342
QY 412 ACCTCATGAGACAGAGTGTCTTCCGCCAGCGCTGATGCAAGCTGGCCTAGCT 471
DB 341 AGATCATGATACAGAGTGTCTTATCCCAAGTATGAGGACAAATGCT-----TGAG 288

```

Oy	472	GCTTCAAGCACAAGAAAGCCAGAGTGGCGAGAGTTCCTTCAAGCGATGTGAACGCTC	533
Db	287	GCTTCAAAACACAGAAATTTCCGTACTCGAAGAGGCATCTGTCTTGCCTTATAGCAACAC	228
Oy	532	TTCAATGATACGGCACCACAGACTTATGTTTGGCTATATATACACACAGTTTGTGCAC	591
Db	227	TCATATGCTCTGGAGGCACAGACTTTTAACACTAAGCAAGATTGTGCACATATATATGCCACT	168
Oy	592	TTCTCGGAGATCCCAAGTTAATATGTAGAGGAGCGCCATCCAAACGCTAATGAAATCT	651
Db	167	TACTTGGAGATCCAAACAGCCAGAGTTCACAGATGACATAAACAGCTTATGTGAATAATT	108
Oy	652	ACAAGCATGTAGGGGATTCGATTCGCGCCACAGACTCCGCT	689
Db	107	ACAGACATGTAGAGGAACGTGTATAGGACAGATCTCACT	70

RESULT 12
 US-10-437-963-83732
 Sequence 83732, Application US/104379663
 Publication NO. US20040123343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boukharov, Andrey A.
 APPLICANT: Bardazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221) B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 83732
 LENGTH: 4171
 TYPE: DNA
 ORGANISM: *Oryza sativa*
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_83034C.1
 US-10-437-963-83732

Query Match	1.1%	Score	54.4	DB	7	Length	4171
Similarity	49.6%	Pred. No.	0.00022				
Best Local							
Matches	226	Conservative	0	Mismatches	221	Indels	9
						Gaps	3

QY	240	ACGGGAGGCACTTTAAAGTTGCACAAAAGTCCCTGAGGGGTTCTGTGGAGCTAAATAAG	229
Db	13	AMGAGCGCACTTCCGCGTCGCGCCAGGGCGGCTCTGACAGGCGCTCTCCGCGCGCGGTG	72
QY	300	CGATTGGGAGCGATTTTAATGCATACACGGCTACCGTTCTGCGCAATGTGATGATCGG	359
Db	73	CTGCGCGGGGACACACTTCAAGATCCACCTCAACGCGCTCTGTGCGCGCGCGGTGACCG	132
QY	360	CTGGGAGACGACGAGGACACAGTCCGCGAAGAGCGCAACTTCTGTGCGGACCTCATG	419
Db	133	CTGGGCGAGCGGACAGACGCGCTCGAGACGCGCGACGAGCTGTGTTACACCTCATG	192
QY	420	GAGCACAAGTGTCTTCCGCCCAAGGCGCTGATTCACAAGCTGTGGCACTAGCTGTTCAAG	479
Db	193	GAGGT-----TTCTATACCGACAATCTATTGTTGAAGGCTGGAAAGTTATGCA-TGGACT	246
QY	480	CACAAAGACGCCAAGGTGCGCGAGGAGTTCCTTCAGACGATTTGGAACGCTCTTCATGAG	539
Db	247	CACAAAGCTGAGAGGTGCGGGAAATAATTGTAGCACTGTTCGAACAAGCTGTTGGGCTTC	306
QY	540	TACGGCACCCAGCAGCTTAGTGT---CGCGTCTAATATACCAACAGTTGTGCACTTCTC	596
Db	307	TTTGGCTTACAGAGAGTCCCTTACAGGAGGTGTGCTTACCTGTCTCGCAATTGCTG	366
QY	597	GGAGATCCCACTTTAATGTGAGGAGGCGGCATCCAAACGTAAGTGAATCTACAG	656

Db			367	AACGATTTCATCAAAGTGTTAGAACCGCTGCATCTCATGTATTGAGGAGATGTATAGA	426
Oy			657	CATGTAGGGGATCGATTGCGGCCAGACCTCCGTCGC	692
Db			427	AACATGGATCTCAATTCCATCAGAGATTGCAGCCC	462

```

RESULT 13
US-10-437-963-89872
; Sequence 89872, Application US/10437963
; Publication NO. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 89872
; LENGTH: 4420
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530.88591C.1
US-10-437-963-89872

```

Query Match	1.0%	Score	53.2	DB	7	Length	4420
Best Local Similarity	46.6%	Pred. NO.	0.0053				
Matches	246	Conservative	0	Mismatches	273	Indels	9
						Gaps	2

200 GGACATGGAGTATTCCTTATTTAGACGTTTATGTCATGCGCTGACGGGACGACCTTTAAGAT 259
 117 GGAGGTACGGGCGCTGTGGACACCTGCATGGATCTGATACGGACGCCAATTTCCGGGT 176
 260 TGCACAAAAGTCCCTGGAGGCGTTCTCGAGAGCTAATTAAGCATTTGGGACGCATTTTAA 319
 177 CGCGCAGGGCGCGCTGGCAGGGCGCTCTCGCCCGCCCGCTGTCGCCGGGAGCACATTTAA 236
 320 TGCATPACGCGCTACCGTTCTGCGACATGTGATCGATGCGCTGGGAGACGACGAGGACAC 379
 237 GATCCACCTGTAACGCGCTCGTCCCGCGCGCCGTGCAGCCCTCGGGAGCGGCAAGCAGCC 296
 380 AGTCCGGAGAAAGGCGCACTTCTGTCGCGGACCTCATATGAGACACAGATGCTTCCGCC 439
 297 CGTCCGAGGCGCCGCGCGGACGCTCTCATACCTCATGGAG-----GTTTCTTGCC 350
 440 CCAGGCGCTGATGCACAAGCTGCGCATTAGCTTCTCAAGCACAGAAAGCGCAAGTGGC 499
 351 GACATCATATTTAAAGAGCGTGGAAAGTTATGCAATGAGCTCACAAAGATTTGAGGGTCCG 410
 500 CGAGAGTTCTTCAAGACGATTTGACGCTTT---CATGATACGGCACCCAGCAGCT 556
 411 GGAAGAGTTGTACGTACAGTGGCAACCGCGGTGGGCTTTTGTCTTCAACGAGCTCCC 470
 557 TAGATTCGCGCTATATATACCAACGAGTTGTGCACTTCTCGAGGATCCCAAGTTATGT 616
 471 ATTGCACGAGATTTCTTCTTACCTGTCTTGCATTTGATGAATGATTCAAATCAAAGTGT 530
 617 GAGGAGGCGGCCATCCAAACGCTAGTGAATCTACAAGCATGTAGGGAGTCGATTGCG 676
 531 TCGAGATGCTCTATCTATTTGCATTGAGGAGATGTACACATATGGGATCTCAGTTCCA 590
 677 CCCAGACTCGGTGCATGACGATGTTCTTGCTCGCAATTTGGCTAT 724

Db 591 TGAAGATTGACGGCCATACTTCCGTACTGCTAAGGAGAT 638

RESULT 14

US-10-357-930-59105/c
Sequence 59105, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endeshe, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-0078CN
CURRENT APPLICATION NUMBER: US/10/357, 930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785, 276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183, 319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189, 862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207, 454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211, 314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219, 007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255, 281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 59105
LENGTH: 615
TYPE: DNA
ORGANISM: Homo sapiens
US-10-357-930-59105

Query Match 1.0%; Score 52.8; DB 8; Length 615;
Best Local Similarity 50.0%; Pred. No. 0.00018;
Matches 163; Conservative 0; Mismatches 157; Indels 6; Gaps 1;
QY 232 CATGGCTGACGGCGCCACTTTAAGATTGACAAAGTCCCTGGAGCGTTCTGGAGC 291
DB 472 CTGGGTGTAACCTTACGCAATTACAGGTGTTCTGCTGGGATGACATCTGTCGCC 413
QY 292 TAAATAAGCGATTGGGCGCCGATTTTAATGCTACACGGCTACCGTTCTGCCACATGTA 351
DB 412 TGGTGACCCCGGCTGACGATCGGTTCAAGGCGCAGATCGGACAGTGTGCCAAGTCTAA 353
QY 352 TCGATCGGCTGGGAGACAGCAGACAGTCCGGAGAGGCGCAACTTCTGTCGCCG 411
DB 352 TAGACAGACTAGAGATGCTTAAGACTCTGTGAGGAGAGACCAAACTCTGCTGCTAA 293
QY 412 ACCCTGATGAGACAGAGTCTTCCGCCCGCCAGCGCTGATGACAAAGCTGGCCACTA 471
DB 292 AGATCATGATCAACTGCTATCTCCCACTAGTATGGACAGAAATGCT-----TGAG 239
QY 472 GCTTACAGACAAAGACGCAAGTGCGCGGAGAGATTCTTCAACGATTTGAAACGCTC 531
DB 238 GCTTCAAAACAAAGATTTCCGTAAGGAGGATCTGTCTGCTTATAGCAACAC 179
QY 532 TTCATGATGAGGCAAGCCGACGCTT 557
DB 178 TCAATGCGTAAGTCTGGGATGAGCCT 153

RESULT 15

US-10-450-763-6570
Sequence 6570, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:

APPLICANT: Hygeq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450, 763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540, 217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649, 167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 6570
LENGTH: 486
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (119)..(478)
OTHER INFORMATION: 33% homologous to Drosophila melanogaster multiple
OTHER INFORMATION: asteris, accession number AF250842, Smith-Waterman Score=171.
US-10-450-763-6570

Query Match 1.0%; Score 50.6; DB 9; Length 486;
Best Local Similarity 52.6%; Pred. No. 0.00071;
Matches 110; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 216 ATTGACGGTTGATGATCCATGCTGACCGGCGACACTTTAAGATTGACAAAGTCCCTG 275
DB 212 ATGACCTGCTACTCTTTGGGTGTAATCTAGCAATTACAAAGTGTCTGCTGGCATG 271
QY 276 GAGCGCTTCGGAATTAATTAAGCGATTGGGCGACGATTTTAATGCTATACGGCTACC 335
DB 272 GACATCTGCTCCGCCCTGTGACCCCGCTGAGATGCTTCAAGGCGCAGATGCGACA 331
QY 336 GTTGTGCAATGATGATCGCTGCGGAGACAGCAGGACACAGTCCGAGAGAGCG 395
DB 332 GTGCTGCCAAGTCTAATAGACAGCTAGAGATGCTTAAGACTTGTGAGGAGACAGGAC 391
QY 396 CAATCTGCTGGCGGACCTCATGAGCA 424
DB 392 CAATCTGCTGCTAAGATCATGATCA 420

Search completed: November 30, 2005, 01:32:32
Job time : 2501 secs


```

OY 530 AQRALREEREGGG-----GGTGTGAPTRRTVSI 564
Db 329 SVQKSLQTYLKSSGVSALPQSDRSSSSQESLNPFSKSMSTANPSTYAGVSAOSSKA 388
OY 565 GRTPGTLOKPTBSMBSISAVDTAAQRAVRAQYTLYSRQKPLGPNNSQAS----- 617
Db 389 SSLPSSLQ-----KSRSIDIVNAAAGAKAH-HAAGOSVRBRGLGAGALNAGSVASLEDT 441
OY 618 ---MGAASAAGSLPRLNLSNNGGTPATTPGSVTRPRBRAC---VSOQPSRSATSP-- 669
Db 442 SDKLGTASDEGRVAKUSA-----PLAMGNAKADRSKRSSTKYVSOQPSRSRGSFPR 496
OY 670 ---STKLDOYXGIGNYRGATGAIPKKASGIPRSTASRETSPTR---SGGLKKRSNY 723
Db 497 VLTTPALSTVSSGVQRLVNSASA--QKRSKIPRSGGSGREASPERLVSABRSRI.PRDSV 554
OY 724 STGA---GSRRTPERNNPYRSAP-----ARLLAQ 750
Db 555 SGGCSREASRESRSRDTSPVRSFQPLASRHSRTGALYAPEVYAGSGPGYGISQSRRLSS 614
OY 751 SREAEHTGVDGDDQPDVYSGDYNSGGMRMRKLMGRDESDIDISEASVCSSESPDS 810
Db 615 SVSARVAVNTGSD--VEEAVADALKKPARR--YESYGMHSDDANSDASCSSESYS- 670
OY 811 YTRGNKSNYSLSGSHTRLDWSTORAPFDDIETIIQFCASTHSEKDGILSTLOYLADGK 870
Db 671 -----RNGSI-----PTYMRQTEDEVAVLRNCASSNMSERKEGILLQIONLTKNOR 715
OY 871 ELTQOOLKCVLDMFRKMPMDHTK--VYSLFDVTVELLIYHANERSRNCSSCLRLNKK 929
Db 716 TLSRVELKRLCETFRMRPADPRGKXVFSMFEITLVDFOVHDDI--QDMLFVLLYOLTKK 774
OY 930 LGTDILNSMHSKIMWTLOYVHEYFTQOLKELFR1ISDSTQPTTKTRIALRFLTDLA 989
Db 775 MGADILGVSQAKVQKALDVTRESFNDQFNLMFTYDQGTPLSKVVALIKETLLA 834
OY 990 NTYCKSSPPPSD--OSQACERTVYLKLAOLADQKSMELRSQARSCVLVALYNLTQMTLL 104
Db 835 ----KOMD-PGGFINSETRLAVSVITWTTEPKSSDVKAQSVLISLFELNTEFMTL 889
OY 1048 LADLPKYVQDASRSCISHSMRQOSCNNGANSFPSSPLSSSPK-----PLQSPSVG 1100
Db 890 LGALPKTQODATKLIHNLHNR---NTGNGTSSMSGSLTRTPRSPANMWSPLSPINT 945
OY 1101 PPSLQSHHQLSISSTSPRSRQSSVE-----QELLPSSELDIOHNIQKTSSE--IRH 1155
Db 946 SQNTLSPAPFVDTENNMSSEDIYSLSRGVTEA1QNFSPSQODNMEPLRKDXKPDGDSM 1005
OY 1152 CFQGGGYQRLALPNGNGHLYQHNDQGOQDSABLSNSTQSSANTQGN---TRESATM 1207
Db 1006 CGG----PGMSDPRAAGDATTBSSQRLALDNKASLHSMPTHSSPRSDYVPYNSDISPF 1064
OY 1208 RLNDLERERTONAKSPPTDDAKVITVSIINMANGELILASINMSESEVVRVALTTKQDPV 1267
Db 1062 NKSA1LKEMAFDDADQPPDD-----LSLDHSD---LVALLKELUS-----NHNERY 1104
OY 1268 ELLQTS1NLNGIC1G1GNCCELPKNGFRS1MRMLN1LLEAEHTDVYIAGLHYSK1MRGNK 1332
Db 1105 EERK1ALVELMKL1TQESFSVWDEHFKIT1LLLETTLDKEPTJ1ALAL1KVLREILRQOP 1166
OY 1328 MRHNMHFLLEL1LK1IOCYOH--SEKALRDIDSM1PRIAPSLPRLSIN1VAVTAGEF 1386
Db 1165 AR--FKYVAEL1VMK1LEAHKOPHKEVVRSAEASVLAT1SPBQC1KVLCP1I1QTDY 1222
OY 1387 PTNLCAIK1LLEVTGHGSEITDAH1DIVFPN1LARSADDTQSMVRKA1VCIK1LV1G 1446
Db 1223 PIN1LAI1KQOTKV1IRVEKET1LL1PE1MPL1IGYDNSEBSYKKA1C1VFL1AVH1A1VG 1283
OY 1447 BEKYKPLSV1NP1SKVRL1N1Y1IEKORNC1SGSGSTKNSSAAS 1490
Db 1283 DE-LKPH1SLQ1LGS1KKMLN1L1Y--KRAQ1SGSGADPT1DVSSQS 1324

```

```

RESULT 2
US-10-104-047-2297
; Sequence 2297, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20030236392A1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104, 047
; CURRENT FILING DATE: 2002-03-25
; PRIORITY APPLICATION NUMBER:
; PRIORITY FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2297
; LENGTH: 916
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-2297

```

	Query Match	16.08%	Score 1214.5	DB 4	Length 916
	Best Local Similarity	32.88%	Pred. No. 1.5e-72		
	Matches 321	Conservative 180	Mismatches 322	Indels 157	Gaps 29
QY	9	LDGFIQMPKADMEVKVQLAEDLVTFISDDTNVIC----	TDWGFLIDGL-MPWLTSGHF	63	
DB	5	MSCCLAQVLQGDVGKRIQVGDELIDYFSDKXSADLEHDDTMRLKLDGLATSWNSNNY		64	
QY	64	KIAQKSLAEPSBELIKRIGSDPNATATVLPVIVIRLDGDSPTVEYKQALLRDLMEHVL		123	
DB	65	KVLLIGMDIDISALVTRLDPRKQAIIGVLPBLIRLDGADKDSVREODOTLLKTKMD-OAA		123	
QY	124	PPQALIKPLATSCPKHNAKAREEFLOTIVAALEHYGTQOLSVMVYIPPVCALLGDPTVN		183	
DB	124	NPQVYMRMLGG-FKHKNFRREBICICLCIATLTAASAGQITLTSKIYVHICNLIGDPSNQ		182	
QY	184	VREAAIOTVEIYKGVDRLEPRDLRMDDVPAKILAEQKFDQYKQEGLLPSALKNTN		243	
DB	183	VRDAIINSIVSEIYHVGEPVRADLSK-KGLPQSRILNVIPTFKEDEVQSGNNIOSA----		236	
QY	244	GNNGVLEADNIGLREPRTRMIKRLPHSNVSSLRPKPNVNDVYG-----		288	
DB	237	NDKNFDEDSVD-GNRRPS-----SASTSSKAPSSRRNVGMGTRRLGSSTLSGK		286	
QY	289	-----DAGATWMSFESSFEVVPOLNIFPAKODMDIYKQVLVIISDKNADMEKRVDAK		342	
DB	287	SSAKKGAAGVDEDFIKAPDVPVVOIYSSRDLEBSINKRELISDDKHMEQGVNMLK		346	
QY	343	KIRALLILSYHTOPQVAVOLKELTSPVDILKEELBSQVIREACTITAYNSKTLRANKID		402	
DB	347	KIRSLLLAGAAEYDNFPQ-HILRLDGA-F-KLSADRLSQVREACTITGLHSLVGNFED		404	
QY	403	AFCSISIEHLINLIQNSAKVIAASSTALKTIKYTAAPKLKLYITDPLNKSXDIST		462	
DB	405	HGAELIMPTLFNLI PNSAKIMATISGVAAVRLIRHTHILPRILPYITSNCT-SKSVAVARR		463	
QY	463	LCEIMAVLLFEEBQKALERNATVLRDYLKKSIGDADCARRSHRYAVAFRRHPBELADQ		522	
DB	464	CFEELDLILDMQVHSLERHISVLAETIKGIIHADSEARLEAKKVCWGFHSFBRABEH		523	
QY	523	IYGLTIDIAQALEREEREGGGGGGTGTGTGTAPETRTYSIRIGTPTGLQKPTPMKSTIS		582	
DB	524	LYHLESSYQKALOSHUK-----NSDSIVSLPQSDRSS--SSEBLNPLRSKSRPT		574	
QY	583	AVDTPAAGQRAKVRQYLLYSRQKRLPDPNNNSQASMTGAAAGSLPRLRLNSGGT--		640	
DB	575	GSTTSRSASTVSKSVSTTGSLSR-----SRSDIDVNAASA--KSKVSSSGCTTFPS		624	
QY	641	-----ATTGGSV-----TPRRGRAG--V		657	
DB	625	SAALLPQGSYASLESRRMRDEMEYIIGDSGRIRRRRSGSGSATVASTPDQRGKSRKVV		684	
QY	658	SQSGPGRSRTSPSTKLDDYGGI-GNYRG-ATGALIPKASGLPRSTASRSTSPTRNSG		715	


```

Db 263 FGGDT-----DITEKPEPVPVHSEKELREBEKIAATLVPEK--DMSVRIANMORI 312
Qy 345 RALLISYHTPOPOFVAVOLKELISFVILKEEL---RSOVLREACITIAVMSKTLRNL 401
Db 313 EALVYGADIDPSPFIML-LKOL-----VPLSTQSDRRSSIVKQCHILANLVSKELLDLF 367
Qy 402 DAFCSHLEHLINLIONSAKYIASASTALAKYITKTAAPKLKITYTTTLMNOSKIDRS 461
Db 368 EPCAEFLFPMLEFKLVITVLVIAESADCIKITILNCKISIRILPRIADTAKNDSAVARA 427
Qy 462 TLCEMLVLLFEEM-OTKALERNATVLRPLKKSIGDA-----DGDARHRSYVAMARR 514
Db 428 RSCETALILLEYWADAPETORSADLYEDLIKCCVADANSDTINDDEGVHCKRSPSIRE 487
Qy 515 HPELADIDYGLDIAOALEREREGGGGTGTGTAPETRRVIRIGRTGTLQKP 574
Db 488 RF-----VQPSRSLSH-----ASGTSALGYTS----- 510
Qy 575 TPRSRSIAVD-TAAOAPAKYRAQTYLXSRQKPLGPNNSQASMTGAAGSLPRPLN 633
Db 511 -----AIYAMDXTAISDSFSSMTLRLSQKTVG--RSSERSLESVLNSSKEKVAID 563
Qy 634 -----SNSGCTPATPGSVTPRPRGAGVSQSGPSR-----STSPSTKLDOYGGIGNY 684
Db 564 HSLONSLALLSSVPSIT-----NASANNGSRLLSEMTTQGTREKRSPLGNIS 615
Qy 685 RGATGAI-----PKKASGIPRSTASRETSPTSRGGGLMKRSMVSTGASRRTPERNRNV 740
Db 616 SESMTSLSPFRPSLERPQSGRNDESDIRS-----TRRFPOQNVVD 660
Qy 741 PSAPARLLAQSREAEHTLGVGDGDPDYVSGDYWRSGMGRKLMGDESDIDSEAS 800
Db 661 -----MPYRDALHR-----DSHNNHVP-NFCRP--LRLKQVMSR-----AS 693
Qy 801 VCSERSFSSYTR-GNKSNY-----SLG-----SGHTLMDSTOAPRPDIETIIQFAS 849
Db 694 ASIHNSFDSQVQSGDVGITDALASLSDALSEGLSPSSDMVVRSAFEFRMLIQ----- 749
Qy 850 THWSEKDLISLTOYLADGKELTQOOLKCVLDMRKMFMTHTKVSFLDYVTELLV 909
Db 750 -----QCGQIGIEIT-----QNFEMVKLFFRLDDBPHKVAQAASSTLAELIPA 794
Qy 910 HANETSRNGSSCLTRLFNKLTDLINMHSKIMKTLOVHEHYFTQOLKELFRITIDS 969
Db 795 -CKRPFESYVERILPYVFSRL-IDPKELVKKFCSSTLDVVGRTYADMLLALVRSI-DE 851
Qy 970 TOTPTTKRIALIRFLDLANTYCSSDPSPQSGACERTVLKLAQLAADKSMELRSQA 1029
Db 852 QKSP--KAKLAVLEFANKSFSKTYVDSGYSN-SGFLKMLSKLAPL-VHEKNAKLKES 907
Qy 1030 RSCIVALT-----NLNTPQMTLLADL-----PKVQDASARCSHMRROQ 1072
Db 908 ISGIISVSHFDSTAVLNFIILSVBEQNLRLAKOYTPRIEVLVLYLOSKORRPPX 967
Qy 1073 S---CNSGANSBPSSPLSSSPKPLQSPSVCPFAS---LOSHHQLSISTSPSRSS 1125
Db 968 SYDQADYCTSSBEDGYALASKSPF---GRYSSSSLDAGGKXWNSVOESTPRNAPMA 1022
Qy 1126 VEOELLESELDIONIKTSEEIRHCFGGQYQTLAANGFNGLQYHDQGOQDSCASLS 1185
Db 1023 -----RTTSDMSIDH-----TSOSIELDTGSEVLT-----RSRE 1052
Qy 1186 SNSKTQSSANTQOS--NTPES-----ATWRDL-----NLERERTONAK 1222
Db 1053 SKNNSTSLVETARSVPNYPEKTDAPLDBETAISTRLDLSHRAADGNAVAGSTAEENVQ 1112
Qy 1223 SPTDADAKITYTSIMANGELILASNLMB-SEVAVALTITDOVEVLLQTSINTLGICI 1281
Db 1113 EGDIAVVKLSIKITLHADNELSIPQLHQISNGTEVSLERKEBALQOLVKASVDN----- 1167
Qy 1282 KGKNCENLNGKPRSRIMRLNILEAHHTDVVYAGLHVLSKIM--RSNMGRRHNMHFEELI 1339
Db 1168 ---DISIAKTFNOLITLAVLEVLDSDSSTREIALSLVAEMLNOSGAMBES---IETV 1220

```

```

Qy 1340 LKTIQCYOH-----SKEALRIDSMTIPRLAPSLPLDLISINYNPVATGEPFNLCALIK 1394
Db 1221 LEKLIHTKXMKVAKISNEANQCNVLAKYDPF-----RCALVVPFLVSSDDEKTLVVCIN 1276
Qy 1395 ILEEVTEHSGEITDAHLDVFPPLNARSADTOSMVKAAVFCIVKLYFVLGEEKVPKL 1454
Db 1277 SLTLVGRLSSEELMANOLPFLPALPFAFSNOSPDAKIVVFLVDIYIMLGKAFV-PYL 1335
Qy 1455 SVLNPSKYRLNLYIEKORNCISG 1478
Db 1336 EGNLSTQRLVLTITVIANRISQARSG 1359

RESULT 5
US-10-437-963-186215
Sequence 186215, Application US/10437963
Publication No. US2004012343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437, 963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 186215
LENGTH: 1308
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_83034C.1.pep
US-10-437-963-186215

Query Match      8.2%; Score 624.5; DB 4; Length 1308;
Best Local Similarity 21.2%; Pred. No. 1.5e-32;
Matches 326; Conservative 279; Mismatches 51; Indels 359; Gaps 60;

```

QY	475	QTKLERNAVATLPDITLKSGIDGDCDARRHRYAUAWFRHFPLEADQITGLDIAAORA	534
Db	355	DAPETQRASDIYEDDIKCCVADAMSEVAATATASCRTMTKTMPESSRLPMSFPDAPYRI	414
QY	535	LEREREGGGGGCTGTGTGTABETRTVSRIGTPECTLOK--PTPSMR-----	579
Db	415	INDE-DGG-----LOKRPSPSLREKGVOLSHASSHA	445
QY	580	-----SIGAVDRAAQRAKVRKQVLYSR-----QKKPLG--PNNNSQMSMCA	622
Db	446	SGTHLAGYSTAIVAMDKSAA-----ISSESSLSHLLSOSKKGIRTAERKIEVLSSEK	501
QY	623	ASGSLPREPLNLSNGGTPATTPGVSVPBPBGAGVQSQP-----GSRSTP---STUKRD	675
Db	502	QKVASIEBLKGNSTLDDSLSPSINTNGSRLVDYTNPHVANKERSRSPYLSLSSES	561
QY	676	QYGGIGNYRGATGAIP---KKAAGIPRSTASNETSPTRSOGGLMKRSMYSTGAGSR	731
Db	562	ISGSSLPYARSSGSGSPYGSTMEESNDTWSTRBMQMDHNYLDMTRYADASHNLHNHQ	621
QY	732	TPERNPPYRBPAPAPALLAOSREAEHTLGVGDGPDIYSGDYWMSGCMRMRKLMGRDES	791
Db	622	VPHFQRPRLKQVARSATSASR---HSF---DDG---HISNDM-----	655
QY	792	DDISEASSVCSERSFDDSYTRGNKS--NYSLSG--HTFLDMSIQAPARDIETIIQFCS	849
Db	656	-----SRYTDGPTSTSDALSGLSASDVARATANFQITLL-----	693
QY	850	THMERKDGILSTLOYLADGKELTQOOLCKVLDMFRKAFMDTHTKVSLFLDTVTEILV	909
Db	694	-HQOK-----GIQWQNONEXKMKLFFRLDDPHHKV/AQAASTLADIIPA..739	
QY	910	HANETSRNGSSCCLTRLFNKLGTDLNLSMHSKIMKTLOVHEHYPTQQLKELFRISDS	969
Db	740	-CKQFESYVRIIIPYVFSRL-IDPKELVROPCSTLEIVGRYPTDITLLPALVSL-DE	796
QY	970	TQTPTTKRIALLFLDNLANTYCKSSPFPDQOAGCRVYLKLAQLAADQKSMELR-SQA	1029
Db	797	QRSP--PAKLAIVLEPANKSPFRKYVDSGEGYAN--SGFLMLATSKLAPL--THEKNALKETS	852
QY	1030	RSCVVALY--NINTEPOMTLADLPKVYDASRSCISHMRP-----OSQSCNSGAN	1079
Db	853	ISGIATVYSHRDSATAVNLFIANTLSTEBQNVYRALKQVTRPIEVDLVNYLOSKEKERSPK	912
QY	1080	SPSSSPSSSSSP-----KPIQSPSVGPASLOSHHQLSISSTSPRSQSSVEOELLFSS	1134
Db	913	SYDQVDPCNSEDGDYALTPKSSYAFGR-----SASSLDNAGSKOMMNVGSS	959
QY	1135	-ELDIQHNQIKTS---EIRRHCPGGOYQOTALAPNGNGHLQYHDOGOQDSCASLSNS	1188
Db	960	IFLDI--STGRSSDVSIDNVKQCFKPEAEV-----LATSR	993
QY	1189	KTQSSANT-----TQSNTPESATYMKLDLNERRTQONKASPTDPAK--VITYSIMA	1238
Db	994	ESKNIAITVVEAARSWTDYPEKSDPATIDENSTGTPLEFRGLAVSDGRGAVISTSVEDA	1053
QY	1239	ENGELILASNLMESEVVR-----ALTUTKOQPELLQTSITLN	1276
Db	1054	QEG-----NPLVELSSVKITPHTSNGPSIPOLINHSVSEVTLSDKRELAQOLVTASTN	1108
QY	1277	LGIICIGKGNCELPNKHFPSIRMLNLIIEAHTDVYAGLHVLSKIMRSNK--MRHNMH	1334
Db	1109	-----NDNSIMWKYFENOILTTTILEVDDDDSSIRELSLVLVEMLHNQDPMEESS--	1158
QY	1335	FLIELILKTIQCYOH-----SKEALRQIDSMIPR-----IAPSLPDLISINYNPVAT	1383
Db	1159	-TEIVLEGLHTVTKDQVAKVENSEAQCINNVYLAQIDPRCLAVIYVPLVSDDEKMLVCT	1217
QY	1384	GEFPTNLCAIKILLEVEHSGSEITDAHLDIVFPPLASADDTQOSMRKAAVFCIVKLYF	1443
Db	1218	N-----CLTUTLVGRLSR-----ELMTQOLBPFLPALPAPAFNQNQSDVAKTYVFCLVDIYI	1267
QY	1444	VLGEKVKPKULSVNPSKVRLLNYIEKORNCISG	1478

[illegible]

PRIOR FILING DATE: 2002-03-05
 PRIOR APPLICATION NUMBER: GB 0105401.4
 PRIOR FILING DATE: 2001-05-03
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: PatentIn Ver. 3.2
 SEQ ID NO 4
 LENGTH: 2065
 TYPE: PRT
 ORGANISM: Xenopus laevis
 US-10-469-780-4

Query Match 4.7%; Score 358.5; DB 4; Length 2065;
 Best Local Similarity 19.3%; Pred. No. 2,7e-14;
 Matches 311; Conservative 266; Mismatches 588; Indels 449; Gaps 70;

13 IQMPKADMKVVKQALAEEL-VTFLSDPTNSIVCTDGMFLIDGMPLMGLSHFPAQKSL 71
 602 MQQDSSNMKERLASMEEFQKTVESMERNDIPCOALVKML-AKPGKETFNFQVMQMKH 660
 72 AFSELIKRLGSDFNAYTATVLPVHIDRLGDSRDTREKAQLLRDIMEHRLVPPQALIDK 131
 661 IVA-LIAQKGNFSTKSAVALDGLVDKVDVK--CGSNAKEALSGIAEACTLPMTA--EQ 715
 132 LAISCFPKAKAKYREELQITIVAAHLEYGTQOLSRYRIYIPVYCALLGDPVANNREAIQT 131
 716 VSLAFQKPKQKQSELTNMLSNAIKEFGFTGINVKAFINVKTALATNPALRTSATIL 775
 192 LVEIYKAVGDRLEPRMDVPSAKLAMEOKFDQVQKQGLLPSALKNTNGVGLDE 251
 776 LGWVLYMGAPLKMFE--EKKRA-LLSQIDAEEKKK--GQTPPVISIRSK-HGSRDE 829
 252 ADNIGLEREPTMIKRELSAVSSSLRPKNVNDVTGACAVTWESFESSFEVVPOLNIF 311
 830 GE-----EGEQDEDPADVT-----DILPRDI- 853
 312 HAQMDIYQVAVLITSDKNADEKVDALKIRALLILSYHQPOFVAVQKELSLSFV 371
 854 ----SDKISSDLVSKIEDKN--WKIRKEGLDEVTAII--NEAKFIQPSISELPSALK 902
 372 DILKELRSQVIREACITIAVMSKTLNKLDAFCMSLIEHLINL-----IONSAKIVA 424
 903 GRLANDS--NKILVQOTITLQOUST-----AMSHNIKQVAKLGMITIVLGDSKANVR 954
 425 SASITALKYIITYHAPKLIKITYDTLNOSKSKDIRSTLCELVWVLFEEWQKALEBNAT 484
 955 AALALGLTKSWVDQGMKWLE--GEDLSEELKKNPFLROELL--GWLAEKLPEMRT 1007
 485 VLBDT-----LKKSIDADCDARRHSRYAVMAFRHFPPELADQIYGLDIAAORALERE 538
 1008 VPEDLQLCVEFLVYNCLEDRNGDVAKKQAEALPIFMMAI-----GFEK- 1049
 539 REGGGGGGTGTGCTATPRTTRVRIGRTPOTLOKTPPSMRISAVDTAAORAKVRAQY 598
 1050 -----MSKATSKILKPAK--DQVVALLEKAKA-- 1074
 599 TLVSROKRLPGPNNSNOASMTGAASGL-PRPRLNSNGCTPA-----TTPGSVTP 649
 1075 ---SMPAKPAPPEKASSKQPAVAQAASAPPAASDSGSGSTDYKPPKKTPTGTOAS 1131
 650 RPRGRAGVSOQPSRSTSPSTKRDQYGGIGNYRATGAIPIKAGAGIRSTASSSETS 709
 1132 KAKTQ---SVSSEGTNTLPSN-----TSLTPSKATNTSLKAKPAQTL 1172
 710 PTRSGGGLMKRSMYSTGASRRTERRNPNVPSAPARLLAQSRAEHTLGVGDGQDPYV 769
 1173 PGK-----KAPSKPAKDEB-----DMSGIYI 1195
 770 SGDYMRSGMRGKMLGRDESDDIDEBASSVCSERSFDSSYTRGKNSNYSLSGSHTRLD 829
 1196 I-----VPMGEQKRVKDE-----KALKV-----LK 1215
 830 W--STORAPF-DLLEITIIQFCASTHMSERK-----DGLISLTOYLADGKELTQQ 876

1216 KNFTPRDEYIENQTLQKMSPCIA--RMLODELPHADPQRIKGLAVMTHELSBKGVISC 1274
 877 LKCVLDMEFKMMDTHTKYSLFLDTVTELLIVHANE-----TSRNGSSCLTRLFKLG- 931
 1275 LDVLKMFTRPFDNTSVLMKCLEYLLKFLMLQOEYHNLTEMGG--TSFLFLYMLKVE 1333
 932 -TDLL-NSMSKIMWTQVNHVEYPTQOLKSLFPIISDSTOTPTTKTIALIRFTDLA 989
 1334 PDIIVAKDVRALLTKRCOV---YPA---SKMFVMEGTYSKSKQBAECLEELGCLV 1385
 990 NTYCKSSDPFDSOACERT---VLKLAQIADQKSMELRSQARGLVALYNNTPQMTL 1046
 1386 ESY-----GANVCQPTPAKALKEIAIHGDDDTYRNALNTIYVYVNHGQVVK 1436
 1047 LPLADLPKYQDSARSCISHMRQOSCSNGANSPPSSPSSPPLOSPEVGPASLIQ 1106
 1437 LIGNL---SEKQMSLEIRIKR-----AGKQAAAPAKQVEKFPQVQSAN--ASIL 1484
 1107 SHHQLSTISIPRRR-----QSVQOELLFSSELD-----IQHNI 1142
 1485 RAPPEDMSSKLNQARNNGHTEPESHVRE--FOLDLEIENDNGTVACEMPALYQHL 1542
 1143 OKTSE-----EIRHCFGGQYOTALAPNGFNGLQYHDQGOQDSCASLSNSTOSSAN 1195
 1543 DEIFEPVLIPEKIR-----AVSPHFDMH-----SNTASTIN 1575
 1196 -----TQSNTPSATMRDLNLEBERTQNAKSPETDAKIT--VSINMAENGSL--- 1243
 1576 FYISQVAVDINASIQALQIDEVLQEDKAEAMSGHIDQFLIATFMQRLAYNTHMADE 1635
 1244 -----ILASNL-----MESEVYRVALLITKQDPVBLLOSLNLNIGICIGKNG 1286
 1636 RLKDDIYELYSCTIGNMISLFPOMESLAREASTGVAKD-----LHGLLSLMDARIEBL 1690
 1287 ELPNKFRISMMLNINLE-AEHTDVVIALGVLSKIMSNNKRRHMMHFLBILKIIQ 1345
 1691 EGGQYVRSVNLVVKLEKSDQNTNISALLMLDQSLATASSPFSSELVMMKCLRMIR 1750
 1346 CYQHS-----KEALRDISMTPIRIAPSLPL-----DLSINIVPIYATGEFTNCAIK 1394
 1751 LPEAINNLNLRILLDHNFM-RVLPKEKELQHKSEMWRTLKTLHT-----LCKLK 1803
 1395 -----ILLEVEHNGSETTDHADIVFPNLABSADDTQSVRK---AAVFCIVKYFVL 1445
 1804 GPKINDHLSMIENKHESEL-EAHLLRV---MKHSIDRTSKGDKETEKASCTI----- 1852
 1446 GEEKYK-----PKLSVLNPSKYRLINVYIEKORNCISGSGSSTKNSS 1487
 1853 -BDKVKANVSDPLAEMFPKIGSKENTKEGLAEIYKKKYSDADIKPFLKSS 1905

RESULT 8
 US-10-408-765A-263
 Publication No. US20040101874A1
 GENERAL INFORMATION:
 APPLICANT: Ghosh, Soumitra S.
 APPLICANT: Fahy, Bojin D.
 APPLICANT: Zhang, Bing
 APPLICANT: Gibson, Bradford W.
 APPLICANT: Taylor, Steven W.
 APPLICANT: Glenn, Gary M.
 APPLICANT: Watnick, Dale E.
 TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 FILE REFERENCE: 660088, 465
 CURRENT APPLICATION NUMBER: US/10/408, 765A
 CURRENT FILING DATE: 2003-04-04
 NUMBER OF SEQ ID NOS: 3077
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 263
 LENGTH: 1972
 TYPE: PRT

ORGANISM: Homo sapiens
US-10-408-765A-263

Query Match 4.3%; Score 328; DB 4; Length 1972;
Best Local Similarity 18.7%; Pred. No. 2,9e-12;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

```

QY 61 SHEFKIAQKSLFAFSELIKRLGSDFNATATVLPVHVIDRIGDSRDYVREKAQILLRDLMEH 120
DB 653 TNFOVMQMKLHIVA-LIAQKGNFSKTSQAQVLDGLVDKIGDVK--CGNNAKEMAMTAIAEA 709
QY 121 RVLPPQALIDKLATSCFPHKNAKAREEFLQTIYNALHEVGTQOLSRYVYIPVYCALLGDP 180
DB 710 CMLPMTA--EQVVSMAFQKPKQKQSEETLNMLSNAIKFEGFSGLNKAFISNVKTALAAAT 767
QY 181 TVNVEEAALQTLVEIKYKVGRLRPDLRRM--DDVPASKLAMEQKFDQVQOEGILLPSA 238
DB 768 NPAVNRATAILGLVMTLYVG---PSLRMFEEDEKFA--LLSQIDAFFE--KMGQSPAP 820
QY 239 LKNTNGNGV-GLDEADNIGLRERPTRMKRLPHSAVSSLSLRKPNVDYTGAGAVTMES 297
DB 821 TGISKHSSTGTDEGED-----GDEPDGSDNDVV----- 849
QY 298 FESSFEVQVQNLNFAKMDNDIKQVLIISDKNADMEKVDALKIRALLLSHTQO 357
DB 850 ----DLRPTBL-----SDKITSELVSKIGDN--WKIRKEGLDVAIGI-----NDAK 892
QY 358 FAVAVOLKELSLFDVILKEELRSQVIREACITIAVMSKTLRNKLDAFCWSILEHLINLQ 417
DB 893 FIOPNIGELPTA---LKGRINDS-----NKI-----LVQOTLNTILO 925
QY 418 NSAKVIASATJALKYIIKYTHAPKLLKIYTDTLNOSKSKDIRSTLCMLVLLFEEWQK 477
DB 926 Q---LAVAMGPNIKHQVKNLGIPI-ITVLGDSKNVRAAALATV----- 965
QY 478 ALERNAIVLBDTLKKSIGDADCDARRHSHRAYAFRRHPELADQIYGTDLIAQALER 537
DB 966 ----NAMABOTKKEWLEBEDLSE-----LKKENPFRLQELGMV---AEK----- 1005
QY 538 EREGGGGGGTGTGTAPETRRATVSRIGRTPTLQKPTSMRSISAVDTAAQOR--ALV 594
DB 1006 ----LPTLRST-----PTDLILCVPHLYGCELRNDVAK 1036
QY 595 RAQYTL-----YSRQKPLG---PNSNQASMTGAASGSLPPRLNNSGGTPTAT 643
DB 1037 KAQDALPFRPMHLGYEKMAKATQKLPKTSKDQ-----VLAMLEKAKVNMKAPAP--- 1086
QY 644 PGSVTPRPGRGAGVSGOSQGSNST---SPSTKLBDQYGGIGNYGATGALPKKASGIP 699
DB 1087 PTKATSKPMGSGAPAFKOPASAPAEBCISSSTEPKPD-----PKKA----- 1127
QY 700 RSTASRETSPTRSSGGLMKRSMYSTGAGSRRTPENNNPVRPAPARLLAQSREAHITLG 759
DB 1128 ----KAPGLSSKAKSAQG---KMPSKTS-----LLEDE----- 1154
QY 760 VGDGQPDYV---SGDYMSGGWRMGRKLM-----GRDE--SPDDISEASSVCS---BR 805
DB 1155 --DKSGPITIVPNGKEQRMKBEK--GLKYLKMNFTTTPREVIYEQLTQWSSCAVKWLDG 1211
QY 806 SFDSSTTRKNSKNSYLSGSHTRLDWSTQAPAPDDITTIIOFCASHTWSEKQGLISLTOY 865
DB 1212 MFSHSDFOHNNKA-----LAVMVD-----HLESEKEGVIGC--- 1241
QY 866 LADGKELTGOOLKCVLDMPKQMDHTTKYVSLFLTVTELLLVANE---TSRNGSSC 922
DB 1242 ----LDLILKMLTLRFDTNTSYLKALYKILFILLSEBYHLTENASSF 1290
QY 923 LTRLEFNKLQ--TDLNLSMHSKIMKTLQVVEHYEFTLOLKELEFRILISDSTQPTTKRIA 980
DB 1291 IPIVLVYKVBKDVIRKIDRALINRMCLVY-----PASKMPFIMEGKSKNSKQRAE 1343
QY 981 IIRFULTLANYYCKSSDPPSDOSQACERT--VLKLAQLAADOKSMELRSQARSCVALY 1037

```

```

DB 1344 CLEELGCLVEST-----GMNVCOPTPGALKELIAVHIGDRBNVANAALNTIVTY 1394
QY 1038 NLNTPQMTLLADLPKRYODSARSCISHMRROSOCSNGANSPSSPLSSSPKPLQSP 1097
DB 1395 NVHGQVFLIGNL-----SEKMSMLEERIKR-----SARPSAAPFKQVEEKPQRAQ 1443
QY 1098 SV-----GPFASLQSHHQLSISSTSPRSOSSVEQELLFSELDIGNICT--- 1145
DB 1444 NISSNANMLRKQPADMSKLNQARSMGHPBAQ--MARE--FOLDLEINDNGTARC 1500
QY 1146 --SEELRHCFGGQYDTALAPNGFNGLQYH--DQGQDSCASSKSTQSSANTQOSNTP 1202
DB 1501 EMPELYQKLDLDFEPLVILPEPKIRAVSPHPDMSNTASTNFIISOVASGDINTSQA 1560
QY 1203 ESATRLNLERERTTONAKSPPTDAKYIVSI-----NMAENGELILASNMESEVAV 1257
DB 1561 LTQLFOISLANEASTGVKDLMHG--LITMLDRIEDLEGQOYIRSVNLLVAVLE- 1617
QY 1238 ALTLTKDQ-----VELLQTSLTNLGICIKGNCGLPKKFRSITMRMLNILEAHTDV 1311
DB 1618 ----KSDQTNILSALLVLLQDSL--LATASSPFSSELVWKCLMRVRLLPDTINSINDR 1671
QY 1312 VIAGLHVLSKIMRSKMHNMWHFLELILKTIQCYOHSKEALRPDISMI-----PR 1363
DB 1672 ILDDHIFMKYPPKELK-----QC--KSEPIRLKTLHTLHTLCKLQKPK 1714
QY 1364 IAPSLPLDLSINIYNVVIATGEPPTNLCAKILLEVTEHSGEITDHALDIVPMLARSA 1423
DB 1715 ILDHLTMDNKK-----ESLEAHLCRM-----KHSMDQGSKSD--KETAKA 1757
QY 1424 D--DTQSMVRKAANVCIVKLYFVLG--BEKVPKLSVLANPSKRLNLNYYIE 1470
DB 1758 SRIDAKSSKAKVNDP--LAEIFKTKGSKENTYEGLAELYEKKKYSADADIE 1806

```

RESULT 9

```

US-10-469-780-6
; Sequence 6, Application US/10469780
; Publication No. US2004014724A1
; GENERAL INFORMATION:
; APPLICANT: FAGAN, RICHARD JOSEPH
; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN
; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN
; FILE REFERENCE: 015959-00003
; CURRENT APPLICATION NUMBER: US/10/469,780
; PRIOR FILING DATE: 2004-02-04
; PRIOR APPLICATION NUMBER: PCT/GB02/00941
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: GB 0105401.4
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 6
; LENGTH: 1972
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-469-780-6

```

Query Match 4.3%; Score 328; DB 4; Length 1972;
Best Local Similarity 18.7%; Pred. No. 2,9e-12;
Matches 278; Conservative 253; Mismatches 543; Indels 416; Gaps 68;

```

QY 61 SHEFKIAQKSLFAFSELIKRLGSDFNATATVLPVHVIDRIGDSRDYVREKAQILLRDLMEH 120
DB 653 TNFOVMQMKLHIVA-LIAQKGNFSKTSQAQVLDGLVDKIGDVK--CGNNAKEMAMTAIAEA 709
QY 121 RVLPPQALIDKLATSCFPHKNAKAREEFLQTIYNALHEVGTQOLSRYVYIPVYCALLGDP 180
DB 710 CMLPMTA--EQVVSMAFQKPKQKQSEETLNMLSNAIKFEGFSGLNKAFISNVKTALAAAT 767
QY 181 TVNVEEAALQTLVEIKYKVGRLRPDLRRM--DDVPASKLAMEQKFDQVQOEGILLPSA 238
DB 768 NPAVNRATAILGLVMTLYVG---PSLRMFEEDEKFA--LLSQIDAFFE--KMGQSPAP 820

```

```

QY 239 LKNTNGNGV-GLEADNIGLREPTRMIRKPLHSAVSSSLRPKPNVDYTGDAVWMS 297
D 821 TRGISKHSSTSGDEGED-----GDEPDGSDNVV-----849
QY 298 FESFEVVPQNLIFIAKMDDIYKOVLIISDKADMEKRDALKIRALLISYHPQ 357
D 850 -----DLPRTEL-----SDKITSELVSKIDKN--WKIRKGEDEVAGII-----NDAK 892
QY 358 FVAVOLKELSLSEFVILKEELRSQVIRACITIAVMSKTLRNKLDAFCWSILEHLINLIQ 417
D 893 FIQENIGELPTA-----LGRLANDS-----NKI-----LVQOTLINILQ 925
QY 418 NSAKVIASASTIALKIYTHAPLKIYDTLNOSKDIIRSTLCMLVLLFEEMWOK 477
D 926 Q-----LAVAMGPNIKQHVKNLGIPI-ITVLGDSKNVRAALATV-----965
QY 478 ALERNATVLRDLTKSIGDADCDARHRSRYVMAFRHFPPLADQIYGTLDIAAQRALER 537
D 966 -----NAMAEOGMKEWLEGEDELSEB-----LKENPFLROELLGWL-----AEK-----1005
QY 538 EREGGGGGGTGTGTAPETRRTVSRIGRTPTLOKPPMSRSISAVDTAAOR-----AKV 594
D 1006 -----LPTLRST-----PTDLILCVPHLYSCLEDNRNGDVOK 1036
QY 595 RAQOYTL-----YSROKPLG-----PNNNSQAMTGAAGSLPRPLNSNGCTPAT 643
D 1037 KAQDALPFFMWHLGEKAKAKATGKLKPTSKQ-----VLAMEKAKVMNAPAPAP-----1086
QY 644 PGSTYPRRGRAGVYOSQOPGSRST-----SPSTKLRYDQYIGYRGATGAIPIKASGIP 699
D 1087 PTKYTSKRMGSAKAPKOPASAPAEDCISSSTEKPPD-----PKKA-----1127
QY 700 RSTASSSETSPTRSGGGLMKRSYSTGASRTRPERNNPVPAAPARLLAQRAREHTLG 759
D 1128 -----KAPLSSAKSAQO-----KMSKTS-----LKXDE-----1154
QY 760 VGDDGQPDYV-----SGDYWRSGMGRKLM-----GRDE-SDDIDSEASVCS-----ER 805
D 1155 -----DKSGPIFIYVPOKEORMDEK-GLKVLKMNFTTPREDEYIEQLKTQMSCVAKWLQDE 1211
QY 806 SFDSYTRGNKSNVSLSGSHRLDMSTORAFDDIETIIQCASTHNSERDGLISLTQY 865
D 1212 MFHSDPFHNNKA-----LAVAVD-----HSEKXGVYGC-----1241
QY 866 LADKELTQOOLKCVLDMFRKMFMDTHKYVSLFDVTELTILVANE-----TSNNGSSC 922
D 1242 -----LDLILKMLTLRFPDINTSVLMKALEYIKLFTLLSEBEYHILNENASSF 1290
QY 923 LTRLEFNKLG--TDLLNSMHSKIMKTLQVHEHFFTOLOKELFRIISDSTOTPTTKTRIA 980
D 1291 IPIYLVKAGPEKQDVIIRDVRAILNRMCLVY-----PASKMPFIMEGTSKNSKQRAE 1343
QY 981 ILRLTLTLANTYCCSSDPDSQACERT-----VLKLAOLADQSMELRSQASCLVALY 1037
D 1344 CLEELGLVSEY-----GMNVCOPTPKALKETIAVHIGDRDNVRAALNTIYVY 1394
QY 1038 NUNTPOMTULLADLPKYVQDSARCSISHMRQOSCNUSGANSPPSSPLSSSPKPOSP 1097
D 1395 NVHGDQVFLKIGNL-----SECDMSLEERIKR-----SAKRPSAARIPKOVEKPORAO 1443
QY 1098 SV-----GPRASIOSHHOLISISTSPRSROSSVEOELLFSSSELDIOHNIQKT---1145
D 1444 NISSNANMLRKGPADEMSSKLNQARSMSGHEAAQ--WVRE--FOLDIDEINDNGVIRC 1500
QY 1146 --SEIHCFCGGOVOTALAPNGFNGLQYH--DOGOOSCASLSNSTQSSANTQSNTP 1202
D 1501 EMPELVOHKLDDIEPVLIPPEKIRAVSPHFDVHNSASTINFIISQVASGDINTYIOA 1560
QY 1203 ESATMRLDNLERERTQNAKSPTDKAYTVSI-----NMAENGELILIASLMESEVRY 1257
D 1561 LTOFQJESLARASTVAKDLMHG--LITLMDSRIEDLEEGQOIVRSVLLVVKYLE-1617

```

```

QY 1258 ALTLTQOP-----VELLQTSITNLGICIKGNCGLPNKHFRTSINRMLNIEAERTV 1311
D 1618 -----KSDQTNILSALLVLLQDSL--LATASSPKFSELVWKCLMRWRLLPDTINSINLDR 1671
QY 1312 VINGLHVLSKINSNKGRRNMWHFLELILKTIIOCYQHSKEALRDISMT-----PR 1363
D 1672 ILDDIHFPMKVPEKELK-----QC--KSEFPRLKITLILHTLCKLKGPK 1714
QY 1364 IAPSLPDISINIVNFIATGEPPTNLCAIKILLEVTEHHGSEITDANLDIYEPNLARSA 1423
D 1715 ILDLHLMINDKN-----ESELHAHLGRMM-----KXSDQTSKSD--KETAKGA 1757
QY 1424 D--DTQSWRKAAVFCIVKLYVLG--BEVYKPKLSTLNSKRLANVYE 1470
D 1758 SRIDAKSSAKAVNDF-LAEIFKXIGSKENTKGLAEIYEVKKYSDADIE 1806

RESULT 10
US-10-469-780-2
; Sequence 2, Application US/10469780
; Publication No. US20040147724A1
; GENERAL INFORMATION:
; APPLICANT: FAGAN, RICHARD JOSEPH
; APPLICANT: PHILIPS, CHRISTOPHER BENJAMIN
; TITLE OF INVENTION: NUCLEAR HORMONE RECEPTOR LIGAND BINDING DOMAIN
; FILE REFERENCE: 015959-00003
; CURRENT FILING DATE: 2004-02-04
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: PCT/GB02/00941
; PRIOR FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 2
; LENGTH: 2038
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-469-780-2

Query Match 4.1%; Score 314.5; DB 4; Length 2038;
Best Local Similarity 18.6%; Pred. No. 2.4e-11;
Matches 271; Conservative 238; Mismatches 517; Indels 433; Gaps 65;

QY 61 SHFKIAQSLSEAFSELIKRIGSDPFNAVTATVLPVHVIDRIGDSRDTVREKAOQLRLDMEH 120
D 659 TNGQVQGMHLIYA-LIAQKGNFSKSAQVLDGLVDKIGDYK--CGNAKAKMTAIAZA 715
QY 121 RVLPQALIDKATSCFKNKAKVREBEPLQTVNALHEVGTQOLSRYVYIPVYCALLGDP 180
D 716 CMLPMTA--EYVVSMAFSQKNPKNQOSETLNLMSNAIKERFGSLNKAFTSNVKTALAAT 773
QY 181 TVNVRERAIOTLVEIKYHGDRLRPDLRM--DDVPAKSLAMLEOKFPDQKQEGLLPSA 238
D 774 NPVARTPAITTLGWMILYVG---PSLRMFEDDEKA-LISQIDAEFE--KMGQSGPPAP 826
QY 161 TVNVRERAIOTLVEIKYHGDRLRPDLRM--DDVPAKSLAMLEOKFPDQKQEGLLPSA 238
D 774 NPVARTPAITTLGWMILYVG---PSLRMFEDDEKA-LISQIDAEFE--KMGQSGPPAP 826
QY 239 LKNTNGNGV-GLEADNIGLREPTRMIRKPLHSAVSSSLRPKPNVDYTGDAVWMS 297
D 827 TRGISKHSSTSGDEGED-----GDEPDGSDNVV-----855
QY 298 FESFEVVPQNLIFIAKMDDIYKOVLIISDKADMEKRDALKIRALLISYHPQ 357
D 856 -----DLPRTEL-----SDKITSELVSKIDKN--WKIRKGEDEVAGII-----NDAK 898
QY 358 FVAVOLKELSLSEFVILKEELRSQVIRACITIAVMSKTLRNKLDAFCWSILEHLINLIQ 417
D 899 FIQENIGELPTA-----LGRLANDS-----NKI-----LVQOTLINILQ 921
QY 418 NSAKVIASASTIALKIYTHAPLKIYDTLNOSKDIIRSTLCMLVLLFEEMWOK 477
D 932 Q-----LAVAMGPNIKQHVKNLGIPI-ITVLGDSKNVRAALATV-----971
QY 478 ALERNATVLRDLTKSIGDADCDARHRSRYVMAFRHFPPLADQIYGTLDIAAQRALER 537

```

```

Db 972 -----NAMAQOTMKEMLEGEDLSEB-----:-----LKKENPFILQRLILGML-----AEK-----1011
QY 538 EREGGGGGGTGTGTGTAPETRTVTRIGRTBGTLOKPTPSMRISIAVDTAAQR---AKV 594
Db 1012 -----LPTLRST-----PDTLILCPHILSCUEDRNDVRK 1042
QY 595 RAQYTL-----YSROKPLG-----PNNSNOASMTGAAGSLPRPLYNSSGTPATT 643
Db 1043 KAQDQLPFFPMHILGEXKAKATGKALPTYSKO-----VLAMEKAKVNPAPAP---1092
QY 644 PGSVTPRPRGAGVQSOPGSRST-----SPSTKLDRDYGIGINYRGATGAI1PKKASGP 699
Db 1093 PTKATSKFMGGSAPAKFQDASAPADICISSSTEPKD-----PKKA-----1133
QY 700 RSTASRSTSPRSGGGLMKSMYSTGAGSRTPPRKNNPVRPSAPARILLAOSREAEHTLG 759
Db 1134 -----KAPGLSSKAKSAQG---KMPKPTS-----LKDE-----1160
QY 760 VGDOGQPDYV---SGDYMRSGGMRGRKLM-----GRDE-SDDIDSPASSVCS-----ER 805
Db 1161 --DKSGPLFIVVNPKEQRMKDEK-GLVYLKKNFTTPEDEYIOUKTQWSSCVAKWLQDE 1217
QY 806 SFDSYTRGKNSYNSLSGSHTRLDWSTQRAFPDDEITIFQCASTHWSERDGLSITQY 865
Db 1218 MFHSDFQHNKA-----LAVWD-----HLESEGEYIGC---1247
QY 866 LADGKELTQOOLKCVLDMFRKMFDTHTKVYSLFDVTETELIVHANE---TSHNGSSC 922
Db 1248 -----LDLILKMLTLFRFDTNYSVLMKALEYLKLFPTLSEBEYHLTENASSF 1296
QY 923 LTRLEFNKG---TDLLNSHWSKIKTKLOVYHHEFYPOLOLKELFILISITQPTTKRIA 980
Db 1297 IPLYLVVAKGEBPDVIRKQVRAILNEMCVY-----PASKFPPIEMGTYSKSKQRAE 1349
QY 981 ILRFYTLDTANTYCKSDPFSDQSOACERT---VLKLAOLADOKSMELRSQARSCLVALY 1037
Db 1350 CLEELGCLVESY-----GMVCOQTPCKALKEIVHIGDNDNAVRAALTYITVY 1400
QY 1038 NLNTPQMTLLADLPKYVQDSARSCTSHMRBROSQSCNSGNSPSSPSSPLSSSPKPLQSP 1097
Db 1401 NVHGQVPEKTLIGNL---SEKQMSMLEERIKR-----SAKRPSAARIKQVEKEPQRAQ 1449
QY 1098 SY-----GPRASLOSHHHOLSISSTSFRKROSSVEQELLSSELDIOHNIQKT---1145
Db 1450 NISSNANMLRKGPAREDMSKLNQARSMGHPEDAQ-WYRRE--FOLDDEIENDNGTVRC 1506
QY 1146 --SEIRHCFGGQYOTALAPNGFNGLDHYHQQGODSCASLSNSKQTSQSANNTQSNTPRE 1203
Db 1507 EMPELVQKLDIDIEPVLIPEPKIRAVSPHDDMSHNSASTINTNTISQVSG--DINTSI 1564
QY 1204 SATWRMLDLERERTONAKSPPTDAKVITVBSINMAENELILASNLSESEVVARVALTLTK 1263
Db 1565 QALQOIDEVLAQEDAEAMSGHIOQFLATMQUL---RLYNHMADEK-----LEK 1613
QY 1264 DQVPEL-----LQTSLTNLGICIKGNCELPNKFRS 1295
Db 1614 DEIITKYSCIIIGNMISLFQIESLAREASTGVKJLHMGLITLMDSDRIEDLEBEOQYIRS 1673
QY 1266 IMRMLNLTLE-AEHPTDVYIAGLHVLSKIMRBNKQRHMMHRLLELILKTIIOCYHSGEAL 1354
Db 1674 VNLVAVKYLESDQNNITISALLVLLODSLATASSPK---FSELVM-----KCL 1719
QY 1355 RDIDSMPIRIAPSLPD---LSINIVNPEI-----ATGEFPT-----NICAIKIL 1396
Db 1720 WRMRRLPBDTINSINLDRILIDIHIFMAYVFKEXKXQCKSEFFIRTLKTLHLHTCKLK--1777
QY 1397 LEVTEHSGSEITDAHLDIV 1415
Db 1778 -----GPKILD-HLTWI 1788

```

```

US-10-307-928A-16
; Sequence 16, Application US/10307928A
; Publication No. US20030229016A1
GENERAL INFORMATION:
APPLICANT: Alsodrook, John P.
APPLICANT: Anderson, David W.
APPLICANT: Boldog, Ferenc L.
APPLICANT: Burgess, Catherine E.
APPLICANT: Caterton, Elina
APPLICANT: Edinger, Shlomit R.
APPLICANT: Gorman, Linda
APPLICANT: Guo, Xiaojia (Sasha)
APPLICANT: Ji, Weizhen
APPLICANT: Kekuda, Ramesh
APPLICANT: Li, Li
APPLICANT: Patcurajan, Meera
APPLICANT: Rieger, Daniel K.
APPLICANT: Shenoy, Suresh G.
APPLICANT: Spytek, Kimberly A.
APPLICANT: Voss, Corine A.M.
APPLICANT: Zhong, Wei
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS
TITLE OF INVENTION: THE SAME
FILE REFERENCE: 24102-502D
CURRENT APPLICATION NUMBER: US/10/307,928A
CURRENT FILING DATE: 2002-12-02
PRIOR APPLICATION NUMBER: 60/341,477
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/341,540
PRIOR FILING DATE: 2001-12-17
PRIOR APPLICATION NUMBER: 60/342,592
PRIOR FILING DATE: 2001-12-20
PRIOR APPLICATION NUMBER: 60/344,903
PRIOR FILING DATE: 2001-12-31
PRIOR APPLICATION NUMBER: 60/373,288
PRIOR FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 60/380,981
PRIOR FILING DATE: 2002-05-15
PRIOR APPLICATION NUMBER: 60/381,495
PRIOR FILING DATE: 2002-05-17
PRIOR APPLICATION NUMBER: 60/383,744
PRIOR FILING DATE: 2002-05-28
PRIOR APPLICATION NUMBER: 60/384,024
PRIOR FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: 60/401,788
PRIOR FILING DATE: 2002-08-07
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 53
SOFTWARE: Curaseqlist version 0.1
SEQ ID NO 16
LENGTH: 2040
TYPE: PRP
ORGANISM: Homo sapiens
US-10-307-928A-16

Query Match          4.0%; Score 304.5; DB 4; Length 2040;
Best Local Similarity 18.5%; Pred. No. 1,2e+10;
Matches 271; Conservative 239; Mismatches 516; Indels 441; Gaps 66;

61 SHFKIAQSLEFSLIRKSGDFNAATVTVPHVIDRIGDSDDTYREKQQLLRDMEH 120
:::|||||:|::|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
TFFVMOKMLHVA-LIAQGNFSKTSAGVVLDGLVDKIDGV--CGDNKEAKNTALAEA 709

121 RVLPFOALIDLKLATSCFGHKNAKVREEEPQTIVNALHEYGTOOLSVRYVPYCALTGDP 180
::|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
EQQVSAIFSCKNPKNOSETLMTLSNAIKEGFSSLNKAFLSNVKLTALLAA 767

181 TVNVREAIIQTLVEIYKHVGRLRPDLRRM--DVDPAASKLAMLEQRFDVOYKGGLPSA 238
|||||||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:~
NPAYRTAITITLGWNYLVVG----PSLRVFEDDEKA-LLSQIDAFFE--KMGGQSPPAR 820

239 LRNTNGNV-GUDEADNIIGLERPTRMTRKPLHSAVSSLRKPPNNNDVTGDGAATMES 297

```



```

Db      : : : : :
821 TRGISKSTSGTDEED-----GDEPDGDSNV----- 849
Qy      : : : : :
298 FESFEVVPQNLIFHAKMDDIYKOVLVISDKNADMEKRYDALKKIRALLILSYHTPOQ 357
Db      : : : : :
850 ----DLPRREI-----SDKITSLSVKIGDKN--WKIRKEGLDEVAGII-----NAK 892
Qy      : : : : :
358 FVAVOLKELSLSPVDILKEELRSQVIREACTTIAVMSKTLNKKDADFCSLLEHLINLIQ 417
Db      : : : : :
893 FIQNLIGELPRA--LKGRLNDS-----NKI-----LVQOTLNLIQ 925
Qy      : : : : :
418 NSAVIASASTIALKVIITKHAPKLKIYDTLNOSSKOIRSTLCMLWLFEBWOTK 477
Db      : : : : :
926 Q---LAAMGPNIKQHVKNLGI-P-IITVLGDSKNVNAALATV----- 965
Qy      : : : : :
478 ALERNAFVLRDTLKSIDADCDARHSRYAWMAFRFRPELADQIYGTLDIAARALER 537
Db      : : : : :
966 ----NAMEQOTGMKEMLEGEDELSE-----LAKENPFLRQELLGWL---AEK----- 1005
Qy      : : : : :
538 EREGGGGGGTGTGTGTABETRRIVSRIGRTPTLOKPTPSMRSISAVDTAAQR---AKV 594
Db      : : : : :
1006 ----LPTLRST-----PTDLICVPHLYSCLEDENGVYRK 1036
Qy      : : : : :
595 RAQYTL-----YSRQKPLG---PNNNOASMTGALASGSLPRRLNNSGTPAT 643
Db      : : : : :
1037 KAQDALPEFMHNLGEXKAKATGKLKPTSKQO-----VLAMEKAVNMPAPAP--- 1086
Qy      : : : : :
644 PGSTPRPRGRAGVSOQPGSRST---SPSTKLDYGGIGNYRGATGAI PKKASGIP 699
Db      : : : : :
1087 PTKATSKMGSSAPAKFQAPAPADCCSSSTEPKPD-----PKKA----- 1127
Qy      : : : : :
700 RSTASSRETPSTRSGGIMKRSMTSGASRRTPERNNPVPAPARLLAOSREAHETLG 759
Db      : : : : :
1128 ----KAPGLSSKAKSAQG---KKMPEKTS-----LKEDE----- 1154
Qy      : : : : :
760 VGDDGQPRVY--SGDYMRSGGMKMGKLM-----GDE--SDDIDSASSVCS-----ER 805
Db      : : : : :
1155 --DKSGPIFIVVNGKEQRMDEK--GLVYLMKNFTTPPEDEIOLKTYMSSCVAKMLODE 1211
Qy      : : : : :
806 SPDSYTRGNKSNVSLSGSHTRLDWSTQARFDEITIIQFCASHTWSEKXGLISLTQY 865
Db      : : : : :
1212 MFHSDFOHMKK-----LAVVND-----HLESEKGVIGC--- 1241
Qy      : : : : :
866 LADGKELTQOLKCVLDMFRKFMMDTHTKVYSLPLDYTELILVHANE--TSRNGSSC 922
Db      : : : : :
1242 ----LDLILKWLTRFPEDTNSVLMKALEVILKLFTELSEEHYTEHNASSF 1290
Qy      : : : : :
923 LTRLPNKLG--TDLNMSHSMKIWTLQVYVHEFPLQOLKELFRILISSTQTPPTTKRIA 980
Db      : : : : :
1291 IPYLVVKKGEKPDVIRKDVRAILNNMCVY-----PASKMFPETMGTSKNSKORAE 1343
Qy      : : : : :
981 ILRFLDTLANTYCSDPFPSDOSQACERT---VLKLAOLADOKSMELRSOARSGCLVLY 1037
Db      : : : : :
1344 CLEELGCLVESEY-----GANVCQPTPGKALKEIVHIGDRNAVNAALNTIYVY 1394
Qy      : : : : :
1038 NLNTPQMTLLADLPKYVQDSARSCISHMRQOSCNSGANSPSSPFLSSSPXPLOSP 1097
Db      : : : : :
1395 NVHGQVQFKLIGNL---SEKDMSWLEBRIKRS-----AKRPAADIKOVEEKPQRAQ 1443
Qy      : : : : :
1098 SV-----GPRASLOS-----HHQOLISSTSPRRROSSVEBELLFSELDIQ 1140
Db      : : : : :
1444 NISSNANMLRGPADMSSKUKIMYRTYRIQARSNSGHPAAQ--VVRRE--FOJLDDEIE 1500
Qy      : : : : :
1441 NIKOT-----SEIRHCGGOYOTALADNGFNGLQVYHDOQODSCASLSSNXTQSSAN 1195
Db      : : : : :
1501 NDNGTVRCMEPEIVQHKDIDIFEPVLIPEPKIRAVSPHFDMDHSTATINFIISQVANS 1560
Qy      : : : : :
1196 TTQSTPESATMRLDNLERERTTONAKSPTDKAYIVYSIMANGELILASNLMESEVV 1255
Db      : : : : :
1561 --DINTSIOALTQIDEVLROEDKEAMSGHIDQFIATFMQD---RLIYVTHMADEK-- 1612
Qy      : : : : :
1256 RVALTLTMDQVEL-----LQTSNLNLGICIKGNCE 1287

```

```

Db      : : : : :
1613 ----LEKDEIILKLYSCIIIGMISLPOIESLAREASTGVKMDHGHITLMDSRIDLE 1667
Qy      : : : : :
1288 LPNKHFRSIRMNLNILE--AEHTDVYIAGLVLSKIMRSNKRHNNMHFLLEILKTIIOC 1346
Db      : : : : :
1668 EGQOIVRSVNLVYVKVLEKSDQNTIISALLVLLODSILATASSPK---FSELVW----- 1718
Qy      : : : : :
1347 YQSHKALRDIIDSMIRIAPSLPD---LSINIVNPI-----ANGEPT----- 1388
Db      : : : : :
1719 ----KCLMRVRLPPTINSINLDRILLDIHLFMKVFPREKJLQCKSEPPFRITLKTLLH 1773
Qy      : : : : :
1389 NLCAKILLEVTEHSGEITDAHDIY 1415
Db      : : : : :
1774 TLCKLK-----GPKIID-HLTM 1790

RESULT 12
US-10-408-765A-2422
: Sequence 2422, Application US/10408765A
: Publication No. US20040101874A1
: GENERAL INFORMATION:
: APPLICANT: Ghosh, Soumitra S.
: APPLICANT: Fahy, Eoin D.
: APPLICANT: Zhang, Bing
: APPLICANT: Gibson, Bradford W.
: APPLICANT: Taylor, Steven W.
: APPLICANT: Glenn, Gary W.
: APPLICANT: Watnock, Dale E.
: TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
: TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
: FILE REFERENCE: 660088.465
: CURRENT APPLICATION NUMBER: US/10/408,765A
: CURRENT FILING DATE: 2003-04-04
: NUMBER OF SEQ ID NOS: 3077
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2422
: LENGTH: 306
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-408-765A-2422

Query Match      3.7%; Score 281.5; DB 4; Length 306;
Best Local Similarity 33.7%; Pred. No. 2,2e-10;
Matches 66; Conservative 46; Mismatches 81; Indels 5; Gaps 4;

Qy      : : : : :
1290 NKHFRSIRMNLNILEAEHTDVYIAGLVLSKIMRSNKRHNNMHFLLEILKTIIOCQYH 1349
Db      : : : : :
109 DEHKITILLLETLGDXEPTIRALVKVREILRHQPAR--FQVAVELFVMTLEAHKD 166
Qy      : : : : :
1350 -SKALRDISMIRIAPSLPDLISINIVNPIATGSEFPNLCIKILLEVTEHSGEIT 1408
Db      : : : : :
167 PKHENVASAEAAVSLTSSPEOCIKVLCPIIQTADYFINLAIKQOTKIERVSKETL 226
Qy      : : : : :
1409 DAHLDIYFNARSADPTQSMVRKAAVEGIVKYFVLGEEKVXPKLVLNPKYRLNLNVY 1468
Db      : : : : :
227 NLLPEIMPGLIGYDIBESSVRKACVCLVAHNAVIGDE--LKHHLQOLTSKKKLLNLY 285
Qy      : : : : :
1469 IEKORNCISGGSGSTKNSAAS 1490
Db      : : : : :
286 I-KRAQTSGGADPTDVSGQS 306

RESULT 13
US-10-425-115-352121
: Sequence 352121, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovacic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B

```

```

1 CURRENT APPLICATION NUMBER: US/10/425,115
2 CURRENT FILING DATE: 2003-04-28
3 NUMBER OF SEQ ID NOS: 369326
4 SEQ ID NO 352121
5
6 LENGTH: 391
7
8 TYPE: PR1
9
10 ORGANISM: Zea mays
11
12 FEATURE:
13
14 OTHER INFORMATION: Clone ID: MRT4577_8429C.1.pep
15
16 US-10-425-115-352121

```

Query Match	Score	DB	Length
3.6%	276	4	391
3.5%	276	7	5810

Qy	42	IVCIDMGILIGIMPLTSHFKIOKSJEASESELIKRIGSPNATYAVLPVHIDRLD	101
Db	35	LTAAEVTLAVDTCMDLTRDANFRINQGGJQALSVAAVAGDHFKIHLNMLVPAVBERLD	94
Qy	102	SRDYREQAOLLRLDMEHVLPPQALDKLATSCFKHAKVREEFLOTTYNALHEYGT	161
Db	95	DKOPHDAAROLLITIME--VSSPTIIVERAGSVAMTHKMSWRDEFFVTVASAVCLFPAS	152
Qy	162	QOLSV-RYIIPVYCALLDPTNVGEAALQTLVEIYKHGBDLRDLRMDVDPAKCLAM	220
Db	153	TEISLQRLVLSPLVLOLMDNSOSVBDALISCEIEMTKHGSQFHEELOR-HNLPSYMLXE	211
Qy	221	LEOKFDQVQEGILLPSALKNTNGNGVGLDEADNIGLREPRMIRPLHSAVSSLR--	278
Db	212	INSRLDKLEPK-----VSSDPTMOYKAVESKGNAGMEV	246
Qy	279	---PKPVNDV---TGDAG-----AVTNESEFSEFEVBPOLNI FPAKMDDIY	320
Db	247	GSCTPGSGVNDMENESGSKGRPLGRDSTKASRKKAAMSSSQSTNYLSRIH-----DIO	301
Qy	321	KQVLYISDKNADEKRYDALKKI--RALLISYHQPPAVQUL-----KELSLSF	370
Db	302	IARLKQSEBEKELKQONIEFMJEVELEKLEFOSKKIEMQOTKLMEDRIRKDEELNKMYA	361
Qy	371	--VDILKEELRSOVIReactITAYMSK	395
Db	362	MDMDALPEELRAVYIARRKKGLIDYFIK	388

Query Match	3.6%;	Score	276;	DB	4;	length	407;
Best Local	Similarly	23.5%;	Pred.	No.	7.9e-10;		
Matches	91;	Conservative	85;	Mismatches	145;	Indels	66;
						Gaps	11;

```

QY 42 IVCDDMGILLDGLMPWLGTSHFHKAQKLEAFSELICKLGDSPFNAYATAVLPHVIVRLGD 101
Db 51 LTAEEVNLAVDTGCDLTRDANFRLAQQGLQALSVAAVAGHFKHLMNLVPAVAERLGD 110
QY 102 SRDVTREKAQOLLRLDMEHRVLPQALIDKATSCFCKHAKAVREBFLOTIYNALHEYGT 161
Db 111 DKQEVHAAAPOLLITLME--VSSPTIIVERGASVAMTHKSMVRBEPRTVASAVOGLFAS 168
QY 162 QOQLSV-BVYIIPVQALLGDPTVNTREALIQTLVEIKYHGDRLDRDLRRMDVVPASKLAM 220
Db 169 TEISLQRLALSPVQLNMWDSNOSVQDAIISCIEEYKIKMGSGOFHELOQ-HNLPSTMLEYE 227
QY 221 LEQKFDVQKOGELLPSALKNTNGNGVGLDEADNIGLERPRMIRKIRPLHSAVSSSLR-- 278
Db 228 INSLDKIEPR-----VASSDPTMQYKAVESGNGMGEV 262
QY 279 ---PKPVRNDV---TGDAG-----AVTMSFESSFEVVPOLNIHAKOMDDIY 320
Db 263 GSGTPSGSVNDMENEESSGKGRPLGRDSTYAKSRKKAMSSSSQSTNYLSRIH----DIO 317
QY 321 KQVAVIISDKRADMEKVDALKKI--RALLLSHTYQGFVAVOL-----KELSLSF- 370
Db 318 IARLKQSEKSELKQONIEFMKEVELKLEFOSKKIEMQOTLYAMEDRIRKOEELNKMYA 377
QY 371 --VDILKEELRSGVIREACTIATVMSK 395
Db 378 MDMDALPEELRAVYIARRKGLIDYFIK 404

```

```

RESULT 15
US-10-437-963-123612
Sequence 123612, Application US/10437963
Publication NO. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with Plant Growth
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ. ID NOS: 204966
SEQ ID NO 123612
LENGTH: 1211
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_2642C.1.pep
US-10-437-963-123612
Query Match 3.5%; Score 264.5; DB 4; Length 1211;
Best Local Similarity 18.7%; Prod. No. 2,6e-08;
Matches 260; Conservative 21; Mismatches 561; Indels 357; Gaps 54

```

Qy	171	PPVCALLGDDPVPNNRREAAIQTCTVITIKYKVGDRLPDLRRMDVPASKLAMEOKPEQVQA	230
Dd	79	PLVAVGRIGDGDPAVRDPAARFLVAKME-----	105
Qy	231	EGLLLPSALXNTN---GVNGVL-----DEADNIGLRERPTMIKRDLHSAVSSSLRPKPVV	283
Dd	106	-----MEMMARBMNGEKTPCISDYQN-----RCMTIQIEPTDTNQAKKSPPKV	149
Qy	284	NDVTGAGAVTMSFESEFVEVYPOLNIFHAKMDDIYKQVLVIISDKNADEKRVDAALK	343
Dd	150	KIVTRDASLLAGADYSAF---PWL-----LRQ	174
Qy	344	IRALLIISTHTPOFOFAVOLKEUSLFVDIIKEELRSOVIREFACTTIYMSTKLANKDA	403

```

Db      175 LVPLI-----TOLL--ORSSVWQACHLNFSLKELLRDEP 211
Qy      404 FCMILLEHLNLIQNSAKVIASTIALKYYIKYTHAPKLIKITYDTLNGSKSDIRSTL 463
Db      212 CAELLIFVLKNVITTHIVIAESSDNCIKEMLRCKARILPKIIEPAKDKSAVLARC 271
Qy      464 CELMVLLEEN-OTKALERNATVLRDLTKSIGA-----DCDARRHRYAWAFRRHF 516
Db      272 CEVALMELMELVDPIDRISVDLYEEFKCIEDATSEMINDEDAFTQORHL----- 323
Qy      517 PELADQYGTIDIAQALEREEREGGGGTGTGTGAPERRRYSIGRPTGLQKTP 576
Db      324 -----SEVERKLMOPQSSSCNST-----EIDKVK-----VD 351
Qy      577 SMRSISAVDTFAAQAQAKVAYOTLYSRORCKPLGNNSNOASMTGAASGLPRPLNSNS 636
Db      352 SGTSFSSVDLOSVO-VKASVQHDMDTKIQPEKTD-----GTTTGS-----SFKDTS 399
Qy      637 GGTPTATPGSVTPPRGRAGV--SOSOPGSHS--TSPSTKLADQYGGIGNYRGATGAI 691
Db      400 TLEKETSPDKGPDPCDGVNLSSCDPSASATQFLEPTSELPSDATV-----VTIV 452
Qy      692 PKRASGIPRSTASRETSPTSSGGGLMKRSMYTGASRRTPERRNNPVRSPAPARLLAQS 751
Db      453 QDAKACRPNTQOLTQOVVLGHPEDLNR-----SPVSLKEG 491
Qy      752 REAEHT-IGVDDDGQPPYVSGDYMRSQGMKRGKLMGRDESDIDSEASSVCSESRFDS 810
Db      492 NSLKNPFIQVNFDS-----SGKGAQKVKHHTSAFN-----KSV 527
Qy      811 YTRGNKSNY-----SLSGHTRLDWSTORAPFDIETIIQFCAS-----849
Db      528 VRKEPRNNYIPNFRRLPLSKQMTNMFYASNSDLDKQLIGEMVSNMDVPSLTEALSL 587
Qy      850 -----TWSEKDCGLISLTQYLA-----GKELTQOLKCVLDMPKMFMDTKVYISFL 900
Db      588 GLKPKSDMMKRYVAFNFLRSLLEQPRGIQEVANQNFQKMFVSRYLADPHHAKIAQAL 647
Qy      901 DTVTELLVHANETSRNGSSCLTRLPNKLGTDLNMSHSMKIWKTLQVHEYPPTQLK 960
Db      648 SLSLEIMPVF-KKFEHYLDKTLPHVFSQ--NDPKESI KOQCSAILTLAGEIYISDILLP 705
Qy      961 ELPRISDSTQPTTKTRIALRFLTDLANT--YCKSDFPDQSQACERTVLAQOLA 1017
Db      706 ALRRTLEEQ--KSPKSKALIEF-----ANSSFVKCTVSSSDNYSSSFLKPMLEKALLF 758
Qy      1018 ADQKSMELRQARSCVALYNTLPOMTLIADLPKYQDSARSCISHMRQOSQCSNG 1077
Db      759 KD-KNSKLKEVSAGFSSISYHDP--ASLLSFLVLSMEQK-----RLRRAIKQLIPT 810
Qy      1078 ANSPSSSPLSSSPK--PLOSFSVG--PFAS--LOSHHOLISSTSPRSROSVEQ 1128
Db      811 IESDLEFLQORGRKPKVPLFDSFVGAYOPYAKGLVIKONKHLHATYOS-----DCLQA 865
Qy      1129 ELFSSELDIQHNIQKTSEIRHCFGGQYOTALAPNGFNGLQYHDQOOD--SCASLSS 1186
Db      866 DQVFDAL--HCLPRISLEVRERBAGKISE--SYDESNG--HDAEMMDKKSSDTRSS 917
Qy      1187 NSKTQSSANTQSTPESATRLDNLERERTONAKSPTDDAKYITVIMAEENGELILA 1246
Db      918 DPPRTFDYNIISDNTVESP--RKEGTDIKREESNESELN-----ITV-----RNRNYIL 965
Qy      1247 SNLSESVVVALTLTKDQPELLQT--SLTNLIGICIGKGNCELPKHFGRSIMGRLNILE 1305
Db      966 -NSGQDGPBSISRTYHQDBMSQMEVEPEINGPSVSK-GICKMLPYCIEIESLQNDLT 1023
Qy      1306 AEHTDVVIAGLHVLISKIMRSNKHNMHFLLELILKXIQCYOHSKEALRD-IDSMIPRI 1364
Db      1024 IHRMSSSL--LEMDDPEESTR-----ELALTLLVEILEKORAMENCIDTLIVKL 1072
Qy      1365 APSLPLDLSINIVPV-----IATGEFTNLCAIKILLEVTE 1401

```

```

Db      1073 LHAATK-DAAKVVNQAHICLTTVYTOFDPULTCLRAIASQLANODEKYLIVINSLSKVI 1131
Qy      1402 HHGSEITDAHLDIYFENLARSADDTQSMVRKAAVFCIYKLYFVLGEBKVRPKLSVINPSK 1461
Db      1132 RLSQDMLMAHLSIFLPALDAPFNHSPYVRKAVVLCIVDTYKLGPAFL-PYLERLDSAQ 1190
Qy      1462 VRLNANYIEK 1471
Db      1191 LQLVITYASR 1200

Search completed: November 23, 2005, 15:16:53
Job time : 212 secs

```